Disabled people and the Web:
User-based measurement of accessibility

André Pimenta Freire

Submitted for the degree of Doctor of Philosophy

University of York

DEPARTMENT OF COMPUTER SCIENCE

September 2012

Abstract

Being able to use websites is an important aspect of every-day life to most people, including disabled people. However, despite the existence of technical guidelines for accessibility for more than a decade, disabled users still find problems using websites. However, our knowledge of what problems people with disabilities are encountering is quite low.

The aim of the work presented in this thesis was to conduct a study that characterises the problems that print-disabled users (blind, partially sighted, dyslexic users) are encountering on the web. This characterisation includes the categorisation of user problems based on how they impact the user. Further, frequency and severity of the main types of problems were analysed to determine what were the most critical problems that are effecting users with print-disabilities.

A secondary goal was to investigate the relationship between user-based measures of accessibility and measures related to technical guidelines, especially the Web Content Accessibility Guidelines (WCAG) 1.0 and 2.0 from the World Wide Web Consortium (W3C). This was done to both identify gaps in the current guidelines, as well understanding where technical guidelines are currently not sufficient for addressing user problems.

The study involved task-based user evaluations of 16 websites by a panel of 64 users, being 32 blind, 19 partially sighted and 13 dyslexics and manual audits of the conformance of websites to WCAG 1.0 and 2.0. The evaluations with print-disabled users yielded 3,012 instances of user problems. The analysis of these problems yielded the following key results.

Navigation problems caused by poor information architecture were critical to all user groups. All print-disabled users struggled with the navigation bars and overall site structure.

Blind users mentioned problems with keyboard accessibility, lack of audio description of videos and problems with form labelling often. However, beyond these seemingly low-level perception and execution problems, there were more complex interaction problems such as users not being informed when error feedback was added dynamically to a page in a location distant from the screen reader.

For partially sighted users, problems with the presentation of text, images and controls were very critical, especially those related to colour contrast and size.

For dyslexic users, problems with language and lack of search features and spelling aids were among the most critical problems.

Comparisons between user problems and WCAG 1.0 and WCAG 2.0 did not show any significant relationship between user-based measures of accessibility and most measures based on technical guidelines. The comparisons of user problems to technical guidelines showed that many user problems were not covered by the guidelines, and that some guidelines were not effective to avoid user problems.

The conclusions reinforced the importance of involving disabled users in the design and evaluation of websites as a key activity to improve web accessibility, and moving away from the technical conformance approach of web accessibility. Many of the problems are too complex to address from the point of view of a simple checklist. Moreover, when proposals are made for new techniques to address known user problems on websites, they must be tested in advance with a set of users to ensure that the problem is actually being addressed. The current status quo of proposing implementations based on expert opinion, or limited user studies, has not yielded solutions to many of the current problems print-disabled users encounter on the web.

Table of Contents

List of Figures

List of Tables

Acknowledgements

The work developed to complete this thesis would not have been possible without the contribution of many people to whom I owe my deepest gratitude.

I would like to thank all the participants who took part in the study for their precious collaboration, in many cases travelling from far afield to come to take part in the evaluations of the websites.

My gratitude also goes to people from a number of organisations who helped greatly with advertisement of the study and recruitment of participants. My thanks to the British Computer Association of the Blind, to Action for Blind People Leeds (now part of the RNIB), the York Blind and Partially Sighted Association and the University of York Student Union.

I would like to express my deepest gratitude to my supervisors Prof. Helen Petrie and Dr. Christopher Power. Their support and guidance throughout the development of this research were fundamental, and I am very grateful for the opportunity to work with them and learn from their vast experience. They are certainly a great inspiration to me and have had a significant influence in my career as a researcher.

Thanks to my examiners Dr. Paul Cairns and Prof. Vicki Hanson for their valuable contributions in the evaluation this work.

I would like to thank the officemates with whom I worked through the years - David Swallow, Lucy Buykx, Anna Bramwell-Dicks, Tara Gilliam, Rasha Salah El-Din, Dina Salah, John Precious, and Nazatul Hamid and other PhD students from the HCI group - Kamal Mohd Othman, Imran Nordin, Cagla Seneler, Frank Soboczenski, Sam Simpson and Sorachai Kamollimsakul. It was a big pleasure working with you all those years! My special thanks go to David Swallow for the discussions about accessibility and constant support in a number of activities involved in this PhD.

I would like to thank all staff members of the Computer Science Department, especially for the operational help during the conduction of evaluations with users and for all the important work they have done that contributed greatly to this work. A very special thank you goes to Filomena Ottaway, former research student administrator, for her support since my application and in the initial years of my PhD.

I would also like to thank for all the friends I met in York during the time I spent there. Thanks for their support and for their friendship, which are very dear to me. A special thanks goes to my friend Alvaro Miyazawa, who started his PhD at the same time as mine, and followed all steps of this journey. Many thanks to my friends Sara , Artur and Lidiane, Miguel, Clarisse and Leo, Luiza, Leandro and family, Lorenzo and Esme, Tom and Tatjana, Frank Zeyda, Carol and David Cross, Simone and Osmar, Richard and Rosa, Ana and Jim, Oleg, Jana and Amir and many others.

My thanks to the members of the University Chaplaincy for spiritual support during my time in York, and special thanks to Fr. Tony Lester for all the support. Many thanks too to my friends of the Catholic Students Society.

I would also like to thank my new work colleagues at the Department of Computer Science of the Federal University of Lavras, for their support in what were very busy months before the completion of this thesis.

Thanks to CNPq – Brazilian Ministry of Science for financial support to develop this PhD, and to the HEFCE for funding part of my tuition fees via the Overseas Research Students Awards Scheme.

A special thanks to my family for their continued support, especially to my parents Teresinha and Leir for their love and care and their efforts to help me succeed. A big thank you to my brother Allyson for being always by my side.

I am eternally grateful to my former MSc supervisor, Prof. Renata Fortes, for her inspiration and guidance in my first steps of my research career, which were fundamental to lead me to do this PhD, and for her friendship and support.

Last, but not least, a very special thank you to my wife Evelise for her continued support and love, and for her patience to endure all the difficulties of being distant from each other during the time of this PhD.

Author’s declaration

I declare that the research work presented in this thesis is original work of my own, unless otherwise indicated in the text. Parts of this thesis have been published in the following papers/book chapters.

1. **Freire, A. P.** ; Petrie, H. ; Power, C. . Deconstructing Web accessibility metrics. *Proceedings of the 3rd Irish Conference on Human-Computer Interaction*. Dublin, Ireland : 2009. v. 1. p. 74-81.

   Some ideas in this paper are used in Chapter 1.

2. Power, C., **Freire, A. P.** and Petrie, H. (2010) 'Integrating Accessibility Evaluation into Web Engineering Processes' in Spiliotopoulos, T., Papadopoulou, P., Martakos, D. and Kouroupetroglou, G., eds., Integrating Usability Engineering for Designing the Web Experience: Methodologies and Principles, IGI Global, 55-77.

   My contribution in this book chapter corresponded to sections describing methods for manual and automated evaluations of web accessibility. Some of this material was used in parts of Chapter 2.

3. **Freire, A. P**. ; Petrie, H. ; Power, C. . Empirical Results from an Evaluation of the Accessibility of Websites by Dyslexic Users. In: 13th IFIP TC13 Conference on Human-Computer Interaction, 2011, Lisbon, Portugal. *Proceedings of the Workshop on Accessible Design in the Digital World 2011*. Aachen, Germany : Sun SITE Central Europe, 2011. v. 792. p. 41-53.

   The results described in this paper were used in parts of Section 4.4 and Section 4.6.6.4 of this thesis.

4. Power, C., **Freire, A. P.,** Petrie, H. and Swallow, D. (2012) 'Guidelines Are Only Half Of The Story: Accessibility Problems Encountered By Blind Users On The Web', in *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, Austin, TX, New York, NY: ACM Press, 433-442.

   My contribution in this paper corresponded to conducting the user evaluations with blind users, and a significant proportion of the data analyses that served as basis to the discussions in the paper. The paper was led by one of my supervisors, who conceived the arguments presented in the paper, with extensive collaboration from myself and the other co-authors. The results from evaluations with blind users and comparisons of user problems with WCAG used in this paper were also used in Section 4.2, Section 5.1.3, Section 5.2.3, Section 5.3.3, Section 5.4.3 and Section 5.5.3 of this thesis.

Chapter 1. Introduction

The Web has become one of the most important and widespread media to provide access to news, services, entertainment and all different kinds of information in people's daily lives. A plethora of daily activities can now be performed using the Web, such as paying taxes, purchasing goods, using banking services, doing online courses and many others. It is clear that improving the Web to make it more used to everyone can have a substantial impact on people's lives.

Disabled people can particularly benefit from having access to services available on the Web, as it provides them with ways to live more independently (Hanson et al. 2009). In order to make websites more inclusive, it is very important to consider that the public may include not only mainstream users, but also users who may have vision, hearing, physical, cognitive disabilities, learning difficulties such as dyslexia or may be from different age groups. The needs of these user groups must be taken into account in the design of websites so that they do not encounter barriers in accessing the Web.

In order to use the Web, many disabled users need to use adaptations in their computers, such as the use of special settings in browsers (e.g. larger fonts, colour changes) or specialised assistive technology. For example, many blind users will use screen reader software that synthesises speech to read content on a web page, while some users who are partially sighted may use screen magnification and changes in the colour scheme. Users may also need alternatives or enhancements to content, such as the provision of textual descriptions of images, audio description of videos, captioning or translation to sign language of audio or simplified versions of text with complex language.

There are several reasons why developers should make their websites accessible to people with disabilities. From a business perspective, it means that websites can reach a wider audience, and hence, expand the range of potential customers. There are approximately 10 million disabled people in the UK, representing 18% of the population (Office for Disability Issues 2011). Besides, several countries have legislation that makes it mandatory to make websites available to everyone, such as the Section 508 of the US Rehabilitation Act (US Government 2011) and the Equality Act in the UK (UK Government 2010). The current ageing of the population is also an important motivation for making websites more accessible (Hanson 2009, Hanson 2001, Kurniawan and Zaphiris 2005), as the prevalence of disabilities is higher with older people (Office for Disability Issues 2011). From the moral perspective, websites should be made accessible because everyone should be entitled to have access to information, products and services, despite of any disabilities.

Despite the importance of making websites accessible, research studies have shown that many websites still present many barriers for disabled users to use e.g. (Coyne and Nielsen 2001, Disability Rights Commission 2004, Leuthold et al. 2008, Petrie and Kheir 2007, Theofanos and Redish 2003). In the largest of those studies, performed by the Disability Rights Commission of Great Britain (2004), it was found that blind users could complete only 53% of the tasks they attempted, showing that accessibility problems can prevent them from having access to a significant amount of information and services on websites. Those findings highlight how critical it is to make websites more accessible and make better websites that disabled users can effectively use.

Technical web accessibility guidelines have been the main resource used to help make websites more accessible. The most well-known are the Web Content Accessibility Guidelines (WCAG), developed by the Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C), with version 1.0 published in 1999 (Chisholm et al. 1999) and version 2.0 in 2008 (Caldwell et al. 2008).

While there have been plenty of studies on the accessibility of websites based on technical guidelines, the number of research studies involving evaluation with disabled users is comparatively small. Those studies are very valuable, as they provide empirical evidence of the context in which accessibility problems occur. Two of the largest studies involving disabled users were developed by Coyne and Nielsen (2001), with 104 disabled users including blind, partially sighted and physically disabled users, and by the Disability Rights Commission (DRC) of Great Britain (2004), with 50 disabled users including blind, partially sighted, deaf and hard of hearing, physically disabled and dyslexic users, both in laboratory and remotely. Other smaller studies on more specific issues with specific user groups have also been performed (Al-Wabil et al. 2007, Leporini and Paternò 2008, Leuthold et al. 2008, Rello et al. 2012, Theofanos and Redish 2003).

There is a lack of empirical evidence to existing web accessibility guidelines, which can be one of the causes of disabled users still finding so many problems in websites. The usability guidelines provided by the US Department of Health and Human Services

(2006) provide ratings with the strength of evidence of each guideline. Such information is not available on the guidelines in WCAG. In fact, some studies have not found evidence of relationship between the evaluation of websites by disabled and conformance to WCAG 1.0 and 2.0 (Disability Rights Commission 2004, Rømen and Svanæs 2008, Rømen and Svanæs 2011).

Building a strong corpus of empirical evidence to support the development of more accessible websites that can be used by disabled users is a clear need (Kelly et al. 2005, Kelly et al. 2007, Petrie and Kheir 2007). Only with strong evidence-based design strategies will web designers be able to produce websites that disabled users will be able to use websites effectively.

Although some large studies, such as the one performed by the DRC (2004), have provided important contributions to building evidence of problems encountered by disabled users on websites, little has been done to build up on the results of such studies. The DRC study revealed several problems encountered by disabled users, but it would be very important to follow leads from this study and deepen the understanding of the nature of problems encountered by disabled users, and the severity of such problems. Many problems encountered in the DRC study with remote evaluations could be further examined with more detailed evaluations performed in laboratory. Besides, many accessibility issues brought with the development and use of new technologies in websites since the DRC study in 2003 need to be explored.

With regard to the relationship between problems encountered by disabled users and technical web accessibility guidelines, the evidence provided by the DRC (2004) and other studies (Rømen and Svanæs 2008, Rømen and Svanæs 2011) have advanced significantly the understanding of this relationship. However, there have been criticisms that there were not enough websites with higher levels of conformance to WCAG in those studies (Brewer 2004). More studies with websites at higher levels of conformance would be able to provide further evidence about the nature of the relationship between problems encountered by disabled users and technical web accessibility guidelines.

The research presented in this thesis aims to expand the body of evidence of problems encountered by disabled users on websites, by performing an empirical study with user-based measurement of the accessibility of websites evaluated by disabled users. The main objective of the study is to further the knowledge of the characteristics of the main problems encountered by print-disabled users on websites.

As a secondary goal, the study also aims to provide further evidence of the nature of the relationship between the problems encountered by print-disabled users and technical web accessibility guidelines. In order to overcome limitations in similar previous studies, this study included websites at different levels of conformance to WCAG 1.0 and WCAG 2.0, and not only websites that did not conform to the guidelines.

In order to achieve advancements in relation to previous related studies, it was necessary to perform a carefully designed study in a larger scale. Instead of several small studies, the research in this thesis consists of one large study that addresses different research questions.

The study focused on users with print disabilities, involving blind, partially sighted and dyslexic users. The restriction on the user groups included in the study was necessary in order to perform a more in-depth analysis of the problems encountered by each group. Results from the DRC (2004) study indicated that these three user groups encountered a wider range of problems than other user groups.

A careful selection of 16 websites was performed for the study. The selection included websites at different conformance levels with WCAG 1.0 and 2.0, selected by means of manual accessibility audits of the home page of hundreds of websites to find enough websites at different conformance levels.

1.1 Research questions and objectives

The main goal of this thesis was to investigate *measures of the accessibility of web sites by means of evaluation by print-disabled users, in order to characterise the main accessibility problems encountered by those users when using websites*. The aim of the characterisation of accessibility problems was to provide researchers and practitioners with a good understanding of the nature of problems encountered by disabled users on websites, how users are affected by those problems and what the technical causes of those problems are.

Based on this goal, the primary research question of this work was:

- **Research Question 1: What are the main characteristics of accessibility problems encountered by print-disabled users when attempting to use websites?**

In order to answer Research Question 1, the following sub-questions were proposed:

- Research Question 1.1: What is the degree to which print-disabled users can complete their tasks on websites?
- Research Question 1.2: How do print-disabled users rate the level of difficulty to perform tasks on websites?
- Research Question 1.3: What are the main types of accessibility problems encountered by print-disabled users on websites and their technical causes?
- Research Question 1.4: What is the frequency of the main types of accessibility problems encountered by print-disabled users on websites?
- Research Question 1.5: What is the severity of the main types of accessibility problems encountered by print-disabled users on websites?

Following the primary goal of this research, a secondary goal was to *investigate the relationship between user-based measures of the accessibility of websites and measures of technical web accessibility*. Having different ways of making a theoretical concept such as web accessibility more concrete and measureable is quite typical in the human and social sciences and is known as operationalisation. As can be seen in

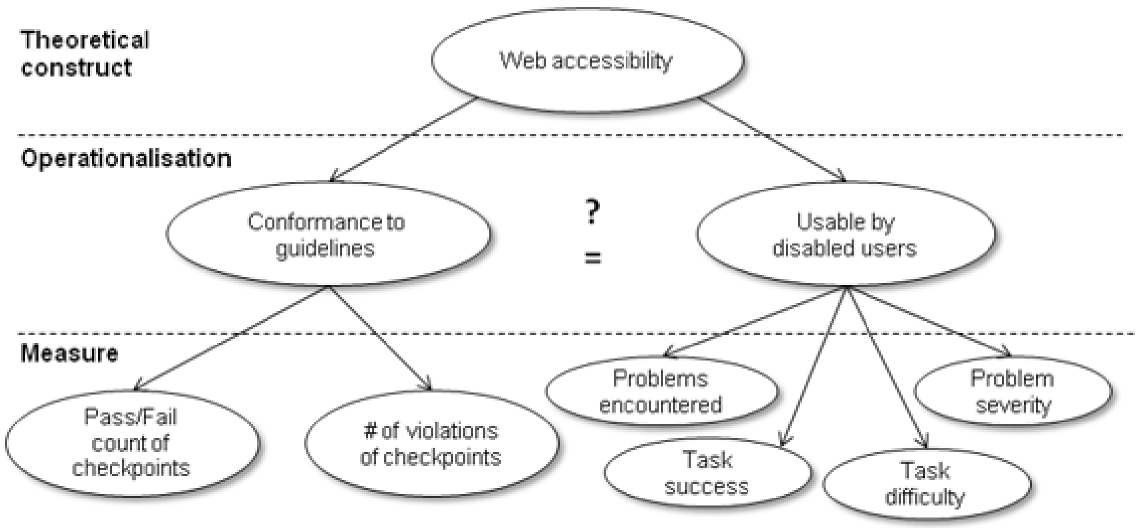


Figure **1.1**, several levels of operationalisation are required to produce a fully concrete measureable construct. Further, there are several routes that can be taken that produce different constructs. However, the relationship between the two routes to measureable constructs depicted in

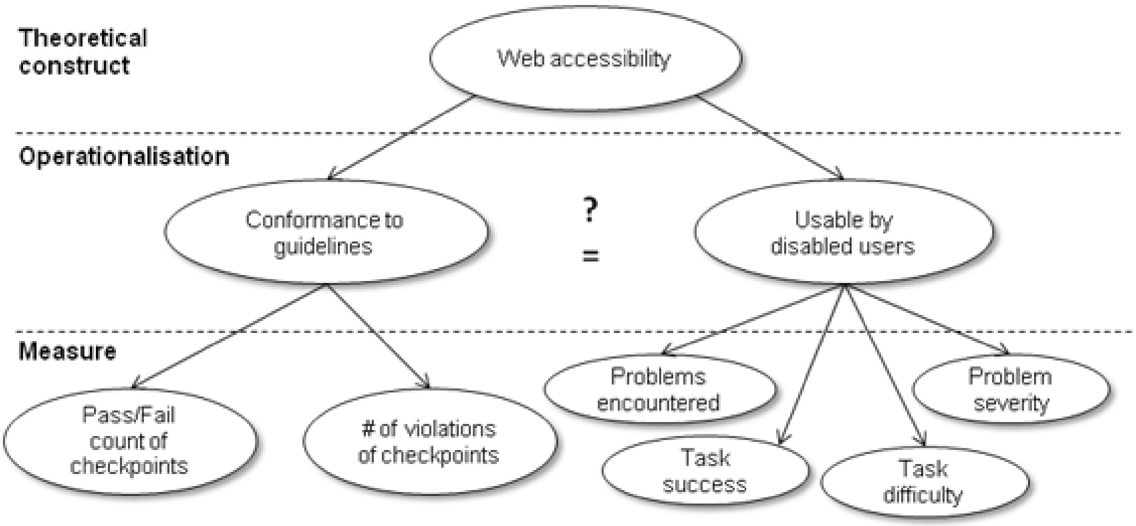


Figure **1.1** has not been investigated. Very few studies have collected data about disabled users' ability to use websites and the conformance of those websites to WCAG. One such study, by the Disability Rights Commission (2004), found no correlation between a range of measures from user evaluation and conformance testing.

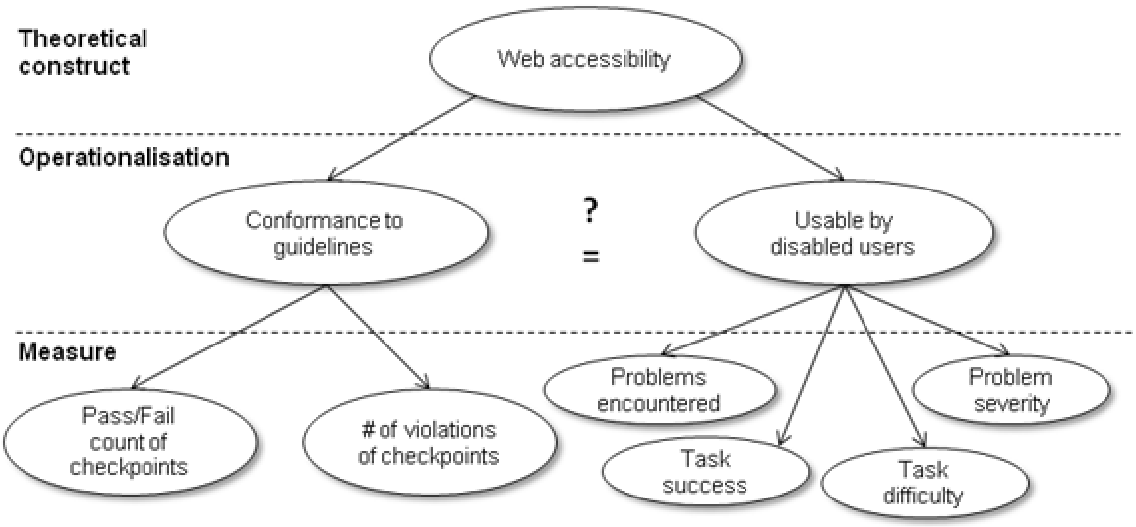


Figure 1.1– Different approaches to the operationalisation and measurement of the theoretical construct of web accessibility

In order to pursue the secondary goal and investigate the two different ways of operationalising the web accessibility construct, the following secondary research question was proposed:

- **Research Question 2: What is the relationship between user-based measures of accessibility of websites and measures of technical web**

accessibility based on the guidelines defined in the Web Content Accessibility Guidelines 1.0 (Chisholm et al. 1999) and 2.0 (Caldwell et al. 2008)?

The following sub-questions were proposed to address Research Question 2:

- Research Question 2.1: Is there any relationship between the number of instances of problems encountered by print-disabled users on websites and the level of conformance to WCAG 1.0 and WCAG 2.0?
- Research Question 2.2: Is there any correlation between the number of instances of problems encountered by print-disabled users on websites and the number of violations of WCAG 1.0 checkpoints/ WCAG 2.0 success criteria?
- Research Question 2.3: What is the coverage of problems encountered by print-disabled users on websites by the guidelines in WCAG 1.0 and WCAG 2.0?
- Research Question 2.4: What is the relationship between the severity levels of problems encountered by print-disabled users and the priority of guidelines in WCAG 1.0 and 2.0 related to those problems?
- Research Question 2.5: Do print-disabled users encounter problems in web pages that conform to guidelines?

1.2 Organisation of the thesis

This thesis is organised in seven chapters and several appendices, containing material used in the evaluations. Chapter 2 presents a detailed review of the literature, examining studies related to user evaluation of web accessibility, evaluation of technical web accessibility and studies that analysed the relationship between the two.

Chapter 3 presents the details of the user study undertaken with print-disabled users. It includes the sampling technique for choosing websites, the methods for accessibility audits and the study design including research participants, materials and the procedures followed during the evaluation session. It also includes a description of how the data from the study was analysed.

Chapter 4 examines the results of the user evaluations and includes information regarding the problems encountered by print-disabled users on the Web. The chapter presents measures from the evaluation of websites by print-disabled users, including measures of task completion, difficulties and number and severity of problems. A detailed description of the frequency and severity of the key types of accessibility problems encountered by disabled users is presented. The full description of all types of accessibility problems encountered, including both users' perspective and technical causes, is presented in details in Appendix D.

Chapter 5 addresses the secondary research question, examining the relationship between problems encountered by print-disabled users on websites and the conformance of websites to WCAG 1.0 and WCAG 2.0. The chapter analyses the relationship between user problems and conformance levels, instances of violations of checkpoints/success criteria, and number of different checkpoints/success criteria violated. It also presents an analysis of the types of user problems that are covered or not by WCAG 1.0 and WCAG 2.0.

Chapter 6 presents the general discussions of how the research questions proposed in this work were addressed in this thesis. Chapter 7 presents a summary of the main contributions of the work in this thesis and presents lines of investigation that could be explored in future work.

Chapter 2. Literature Review

This chapter presents a review of the literature focused on the evaluation of the accessibility of websites. The chapter presents studies that concern evaluation of the accessibility of websites by disabled users, studies with evaluation of the conformance of websites to technical guidelines, and studies that investigate the relationship between user evaluation and technical guidelines. Section 2.1 presents user-centred definitions of accessibility and Section 2.2 presents technical web accessibility guidelines. Section 2.3 presents evaluation of web accessibility with disabled users and evaluation based on technical guidelines. Section 2.4 presents a review of related studies with the evaluation of websites with disabled users, Section 2.5 presents studies that compared evaluation of websites with disabled users with technical guidelines.

2.1 User centred definitions of accessibility

The concept of web accessibility has been related to the issues related to the use of websites and web applications by people with disabilities. However, a clear and comprehensive definition of accessibility has not still been agreed as a result from some confusion in different definitions (Petrie and Kheir 2007, Petrie and Bevan 2009, Yesilada et al. 2012).

Other authors have proposed alternatives to defining accessibility from the user perspective. Shneiderman (Shneiderman 2000, Shneiderman 2003) proposed the term *universal usability* stating that accessibility would be a precursor to usability. Thatcher *et* al. (2003) defined accessibility as being a disjoint subset of problems of people with disabilities from mainstream users. However, results reported by Petrie and Kheir (2007) have shown that there is a common subset of problems affecting both users with disabilities and mainstream users, as well as problems that affect each group separately. Besides, other studies have also shown that there are usability problems that affect mainstream users whose effects can be amplified for users with disabilities (Disability Rights Commission 2004, Harrison and Petrie 2007).

The definition of accessibility provided by the International Standards Organization has made it closer to that of usability. According to the ISO 9241 standard on

Ergonomics of Human System Interaction- Part 11 (International Standards Organization 1998), *usability* is defined as:

> "The extent to which a product [service or environment] can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.".

In this same standard, *effectiveness* is defined as "the accuracy and completeness with which users achieve specified goals"; *efficiency* is defined as the resources expended in relation to the accuracy and completeness with which users achieve those goals; and *satisfaction* is defined as "freedom from discomfort, and positive attitudes towards the use of the product [system, service or environment]".

Part 171 of ISO 9241 (International Standards Organization 2008) on software accessibility, defines accessibility as:

> "the usability of a product, service, environment or facility by people with the widest range of capabilities".

The definition provided by ISO 9241-171 seems to extend the definition of usability of software products to "people with the widest range of capabilities", particularly those with disabilities.

In the context of the work presented in this thesis, emphasis is given to a user-based definition of accessibility. The definition of the term "web accessibility" used in this work is adapted from the definitions from ISO 9241-11 (International Standards Organization 1998) and ISO 9241-171 (International Standards Organization 2008), and used by Petrie and Kheir (2007) as:

> "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".

2.2 Technical accessibility

Technical accessibility is defined by whether or not web content that is implemented on a web page meet criteria that are specified in one or more sets of guidelines (Arrue et al. 2007, Cooper et al. 2012, Henry 2003).

. There are several different guideline sets that have been proposed; however, the most famous and arguably the most important are the Web Content Accessibility

Guidelines (WCAG) from Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C).

Other guidelines and governmental web accessibility policies were also defined by government bodies, such as the Section 508 of the US Rehabilitation Act (US Government 2011) and the Web Accessibility Code of Practice published by the British Standard Institute (British Standards Institute 2010).

The model of accessibility proposed by the WAI is composed of three main sets of guidelines: the Web Content Accessibility Guidelines (WCAG) (Chisholm et al. 1999, Caldwell et al. 2008), the Authoring Tool Accessibility Guidelines (ATAG) (Treviranus et al. 2000) and the User Agent Accessibility Guidelines (UAAG) (Jacobs et al. 2002). This model provides guidelines to be used by developers of web content, developers of authoring tools and user agents (such as web browsers and assistive technologies), expecting that web content in conformant web pages, produced by conformant authoring tools and rendered by conformant user agents would make for websites to be accessible to disabled users.

The Web Content Accessibility Guidelines (WCAG) were developed by the Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C) to encourage and ensure the development of accessible content for the web. Version 1 (WCAG 1.0) was released in 1999 (Chisholm et al. 1999) and Version 2 (WCAG 2.0) in 2008 (Caldwell et al. 2008).

WCAG 1.0 comprises 14 high-level accessibility guidelines, which are broken down into 65 more specific checkpoints. Each checkpoint is assigned a priority level (Priority 1, 2 and 3). A Web page or resource must satisfy Priority 1 (P1) checkpoints otherwise, according to WCAG 1.0: 'one or more groups [of disabled people] will find it impossible to access information in the document' (Chisholm et al. 1999). If Priority 2 (P2) checkpoints are not satisfied, one or more groups of disabled people will find it difficult to access information in the document. If Priority 3 (P3) checkpoints are not satisfied, one or more groups of disabled people 'will find it somewhat difficult to access information' (Chisholm et al. 1999). If a website passes all P1 checkpoints, it is Level A conformant; if it passes all P1 and P2 checkpoints, it is Level AA conformant; and finally if it passes all P1, P2 and P3 checkpoints, it is Level AAA conformant.

Problems and limitations of WCAG 1.0 were reported in a number of studies, involving difficulties in understanding the guidelines, the interdependency of WCAG on other guidelines, ambiguity, logical flaws, the closed nature (especially with regards to the limitation to W3C technologies) and the complexity of the guidelines (Colwell and Petrie 2001, Donnelly and Magennis 2003, Kelly et al. 2005, Sloan et al. 2006).

WCAG 2.0 starts with four high level principles of web content accessibility: that content must be perceivable; interface components in the content must be operable; content and controls must be understandable; and content should be robust enough to work with current and future user agents (including assistive technologies). Each principle has its associated guidelines, referring to different aspects of accessibility, comprising a total of 12 guidelines. Further, Guidelines under each of these Principles have been rephrased to be solutions to specific user requirements, such as the provision of text alternatives for non-text content (Perceivable, Guideline 1.1). For each Guideline, there are Success Criteria (SC). SCs are testable statements that a developer can use to determine if web content on a web page is accessible. It is against these SC that a website is measured for conformance, with each SC having a priority level, Level A, AA or AAA, relating to conformance levels that are similar to WCAG 1.0.

Meeting all success criteria with a certain priority level is the first of five requirements to achieve a certain level of conformance to WCAG 2.0. The second requirement is that a full page has to be conformant, excluding the possibility to achieve conformance of only parts of pages. The third requirement is that, if a page is part of a process that involves several steps, then all pages in the process have to be conformant. The fourth requirement states that “only accessibility-supported ways of using technologies are relied upon to satisfy the success criteria” (Cooper et al. 2010a), meaning that all technologies used on a web page have some level of accessibility support provided by assistive technologies and user agents. The fifth requirement states that “if technologies are used in a way that is not accessibility supported, or if they are used in a non-conforming way, then they do not block the ability of users to access the rest of the page”.

In order to future-proof WCAG 2.0 during the current times of fast technology evolution, the WAI removed the technical aspects of accessibility from the Guidelines and SC. Technical information regarding how to implement web content with existing web technologies is now provided in separate documents (Cooper et al. 2010a). These documents describe techniques that have been determined by the WCAG Working Group to be “sufficient for meeting the success criterion” (Cooper et al. 2010a). For each SC there can be any number of sufficient techniques for meeting the criteria; however, if a developer can show that they have another implementation that satisfies the criteria, for example through user testing, they need not use one of the WAI approved techniques.

A serious concern about the sufficient techniques is that there is little evidence to support the claims that they are "sufficient". In fact, a study conducted by Power et al. (2011) with 25 visually impaired users showed that different techniques to implement links were not as effective. The destination of links implemented with some of the techniques was only correctly identified by fewer than 50% of the participants in the study.

2.3 Evaluation of web accessibility

The evaluation of the accessibility of websites can be performed using many different methods, some involving real users attempting to perform tasks others that involve experts in accessibility reviewing websites according to principles or guidelines or the use of automatic evaluation tools. This section presents the main aspects of methods to evaluate web accessibility, both involving users and expert-based evaluations.

2.3.1 Conformance evaluation

The evaluation of a web page, website or web application for its conformance to the Web Content Accessibility Guidelines (WCAG) is one type of measure of the accessibility of a website. Accessibility audits by means of conformance evaluation consist of checking the features of a website as to whether they satisfy the conformance criteria that are specified in WCAG. For web engineers who are familiar with accessibility, this is the most common type of evaluation done due to the influence that WCAG has had on the legal and political landscape, given by the requirements to meet WCAG by laws in countries such as Australia, India, the Netherlands and others.

In a *conformance evaluation* a web accessibility expert goes through each guideline checking the features of a website against the criteria of that guideline[1]. It can be undertaken through *conformance tests* conducted via a combination of evaluation with automated testing tools and manual inspections where experts compare web page implementations against guidelines. When such an evaluation is undertaken, some of

---

[1] Checkpoints in WCAG 1.0, Success Criteria in WCAG 2.0.

these criteria, such as the presence or absence of alternative text, can be checked with an automatic tool. In other cases, such as criteria relating to the clarity of the contents of the alternative text, the evaluation can only be conducted using human judgment.

Different methods exist that provide guidance to perform conformance evaluations using both automated evaluation tools and manual inspection, such as the conformance evaluation process defined by the WAI[2] and the process used for accessibility audits applied in the Digital Media Access Group (DMAG) (2002, 2006).

In this section, the different types of tests that can be undertaken in a conformance evaluation are presented, both automated and manual inspection

2.3.1.1 Conformance tests with automated accessibility testing tools

Automated tools can be useful tools to help evaluators check accessibility issues which would be otherwise very tedious for evaluators to check manually. For example, these tools can check the validity of (X)HTML mark-up and the use of style sheets. This can include checking features such as the correct nesting of elements in tables and headers, and proper use of other W3C recommended technologies. This first step helps ensure that a web page contains basic structuring elements that will enable it to be read by assistive technologies. Tools that perform automated checking of checkpoints can be useful in the evaluation of prototypes or initial versions of websites, in order to detect basic accessibility problems early in the development (Petrie and Bevan 2009).

The features available in automated tools can help checking a subset of WCAG in a less time-consuming manner and are heavily used by practitioners (Ivory 2003).

Beyond the technical tests regarding basic mark-up, those tools can check things that can be detected automatically. They can check the presence or absence of features, such as alternative text attributes and headings, or can check values against known pre-defined values, such as values for colour contrast defined in a set of guidelines. The results of all of these tests are usually presented to the user in the form of a report that details problem areas of the web page(s) for the developer. Figure 2.1 shows an example of a report produced by the tool Hera (Benavídez et al. 2006), with a

---

[2] Available online at http://www.w3.org/WAI/eval/conformance.html, last accessed on 04/09/2012

report of an evaluation with WCAG 1.0 indicating problems with checkpoints 9.5 and 10.4.

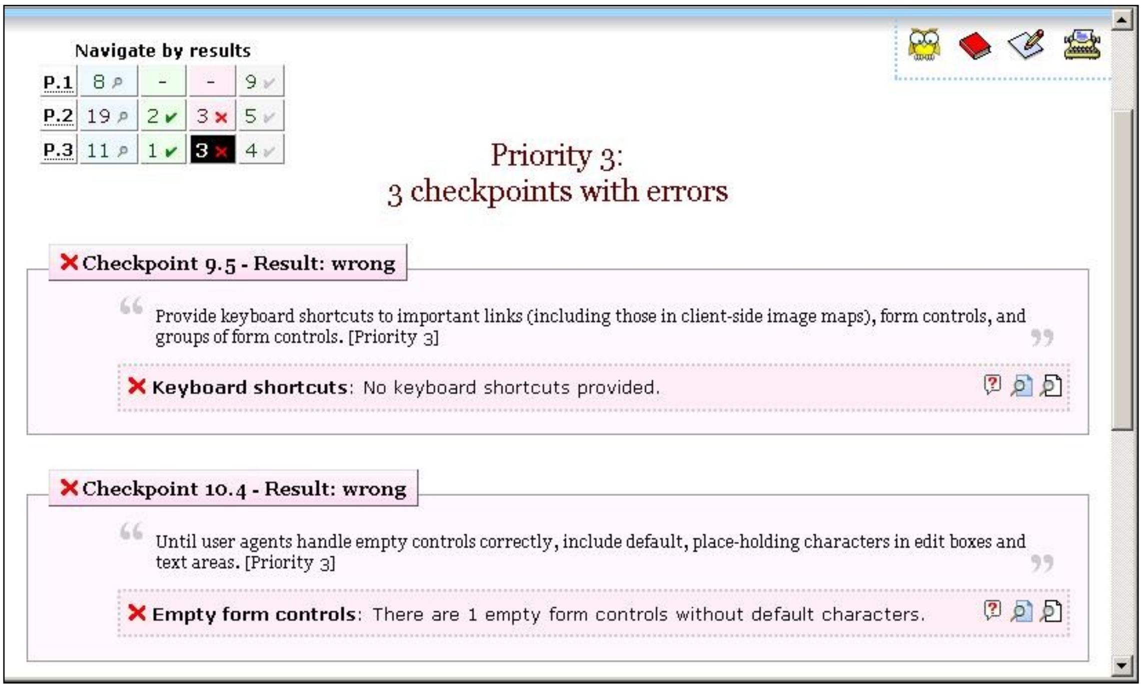


Figure 2.1– Example of report produced by automatic accessibility evaluation tool Hera (Benavídez et al. 2006)

Since the publication of WCAG 1.0 (Chisholm et al. 1999), a number of automatic evaluation tools have been developed to perform tests to check the conformance with the guidelines (Abascal et al. 2004).

One of the most widely used accessibility evaluation tools was Bobby, developed by CAST, then bought by Watchfire and now owned by IBM as the "Rational Policy Tester Accessibility Edition"[3]. Other tools include Wave[4], Hera (Benavídez et al. 2006), Imergo (Mohamad et al. 2004) and many others. However, although WCAG 2.0 was published in 2008, until the moment when this thesis was written, few automatic evaluation tools based on WCAG 2.0 had been made available. The only tools released as stable

[3] Available at http://www-01.ibm.com/software/awdtools/tester/policy/accessibility/index.html

[4] Available at http://wave.webaim.org, last access on 04/09/2012

products available to evaluate with WCAG 2.0 at the time were Total Validator[5], eXaminator[6], A-Checker[7].

However, in context of an evaluation, it is important to highlight that automated evaluation tools are very limited in their capabilities. Although they may help to identify problems that otherwise would very tedious to test, there is only a small number of WCAG guidelines that can be tested automatically. For example, the Unified Web Evaluation Methodology (UWEM) (Web Accessibility Benchmarking Cluster 2007) defines a set of methods and accessibility test cases. From a list of 108 test cases listed at UWEM for the WCAG 1.0 checkpoints, only 26 of the tests (less than 20%) can be checked with an automatic tool. Although many of these automatable tests may help considerably to reduce time and effort spent in evaluation, it is clear that, even for evaluation based on checkpoints review, relying exclusively on automated tools covers only a very limited number of problems users may encounter. As an example, consider the use of text alternatives for images. Although it is possible for an automatic tool to identify whether an image element has an *alt* attribute, the tool cannot identify if the text contained within that attribute describes the image appropriately.

Understanding the outcomes of an automated accessibility evaluation tool is also frequently a burden to evaluators and developers (Choi et al. 2006). Even experienced evaluators very often face problems in comprehending what the error messages mean. Although the so called “warning messages” may help find potential errors in a manual checking, these messages are many times vague and obscure, and end up not clearly showing a good clue where the problem may be, or more importantly, how to repair it.

Finally, there is a question of validity of automatic evaluation tools. The implementation of the checking algorithms varies substantially between different tools, and validation tests for the tools are often not available. This can lead to inaccuracies in checks, such as those found by a study performed by Brajnik (2004), where he identifies reporting errors in various tools.

The relative readiness with which accessibility evaluation results can be obtained with automatic evaluation tools has motivated the use of such tools in a number of research studies of the accessibility of websites. There have been a number of studies in several areas, such as the evaluation of governmental, (Al-Khalifa 2012, Goette et al.

---

[5] Available at http://www.totalvalidator.com/, last access on 04/09/2012

[6] Available at http://examinator.ws/, last access on 04/09/2012

[7] Available at http://achecker.ca, last access on 04/09/2012

2006, Lazar et al. 2010, Evans-Cowley 2005, Paris 2006, Potter 2002, Yuquan 2007) education websites (Espadinha et al. 2011, Hackett and Parmanto 2005, Kane et al. 2007, Kelly 2002), and other cross-section studies (Hackett et al. 2005, Lopes and Carriço 2010, Loiacono and McCoy 2004, Lopes et al. 2010). Many of these studies have pointed out that the lack of accessibility in websites is a serious problem in many sectors. However, the studies cannot identify all problems present in the websites, as tests that cannot be performed by automatic evaluation tools are not included.

2.3.1.2 Conformance evaluations with manual inspection by accessibility experts

Along with tests with automatic tools, accessibility audits based on manual inspection methods by expert evaluators play an important role in the evaluation process of web applications. The use of inspection methods is important to help finding barriers in web resources that cannot be checked automatically. Although they cannot uncover all the problems that users may encounter, these tests are good to find problems early in development.

Manual inspections of accessibility may be performed with the help of other tools to help perform specific tests, such as checking colour contrast, simulating the visualisation of a web page in specific conditions (different colour background, font size, with and without javascript, for example). Manual audits should also involve tests with specific assistive technologies used by people with different disabilities, such as screen readers, screen magnifiers, and using the interface with keyboard only.

Besides the specific automatic evaluation tools to evaluate accessibility guidelines, a number of supporting tools can be used to support evaluators to perform manual accessibility audits of websites. The tools include multi-purpose toolbars that help perform several different tests and are used in internet browsers, such as the Firefox Web Developer Toolbar[8] and the Web Accessibility Tool Bar[9] for Internet Explorer and Opera, developed in a partnership between Vision Australia[10], The Paciello Group[11] and

---

[8] Available at https://addons.mozilla.org/en-US/firefox/addon/web-developer/

[9] Available at http://www.visionaustralia.org.au/ais/toolbar/

[10] Available at http://www.visionaustralia.org/

[11] Available at http://www.paciellogroup.com/

the Web Accessibility Tools Consortium[12]. These tools provide features such as resizing the text of web pages, showing the alternative texts of images, disabling images, highlighting and displaying information about forms.

The ability to use and understand the set of guidelines being used in manual inspection by evaluators are essential aspects of manual inspections. There have been a number of studies have been that indicate many problems with WCAG 1.0 and WCAG 2.0.In one such study, performed by Colwell and Petrie (2001), problems with the navigation in the WCAG 1.0 documents and with the language of the guidelines were found following a study with 12 experienced web developers, who had little knowledge of accessibility. The presence of ambiguities and use of technical jargon in WCAG 1.0 were also pointed out in other studies (Donnelly and Magennis 2003, Kelly et al. 2005, Sloan et al. 2006). In a later study by Petrie et al. (2011), results from interviews with 14 web accessibility evaluators revealed that the line between what could be evaluated automatically and what needed to be evaluated manually was not clear to evaluators.

The level of expertise of the evaluators has a significant impact on the results from manual inspections, and studies have shown that there can be substantial discrepancy in results obtained from different evaluators. A study performed by Yesilada et al. (2009) found significant differences between expert and non-expert evaluators performing manual inspections. In a follow-up study performed by Brajnik et al. (2010), involving 22 expert and 27 non-expert evaluators using WCAG 2.0, it was found that the agreement level of 80% could not be reached for 50% of the success criteria. Further to this, Brajnik et al.'s study also found that 32% of previously known problems were missed by non-expert evaluators. A similar study performed by Alonso et al. (2010) with 25 non-expert novice evaluators also showed problems with the consistency in the evaluations performed by them. The results from these studies all confirm that the experience and understanding of the guidelines by evaluators can have a significant impact on the outcomes of manual inspections.

2.3.2 Other expert inspection methods

Another method based on manual audits is the Barriers Walkthrough method (Brajnik 2006), which was inspired in the use of usability heuristics to perform

---

[12] Available at http://www.wat-c.org/

walkthrough evaluations. The method is based on the concept of detection of *barriers* for users with different types of disabilities.

The method adopts the concept of *barrier* as "any condition that hinders the user's progress towards the achievement of a goal" (Brajnik 2006). The method provides evaluators with a list of possible barriers, which are described according to 1) the types of users and types of disabilities that may be affected, 2) the type of assistive technology being used, 3) the *failure mode* (activity or task that may be impacted by the barrier) and 4) the consequences of the occurrence of the barrier. The list of barriers used with the method is classified according to groups of users separated by types of disabilities.

In two studies (Brajnik 2006, Brajnik 2008), a comparison between a conventional checklist review and the barriers walkthrough method showed the latter to be better in several aspects. The barriers walkthrough was shown to be more *precise* (problems found are more prone to be *true* problems), to lead to a smaller number of reports of *false* problems and to be better to identify more severe problems.

However, according to the second experiment comparing the methods (Brajnik 2008), the barriers walkthrough had low inter-rater reliability, as independent evaluators tend to produce different results. In particular, the barriers list provides a level of understanding to the evaluator as to what each of the barrier means, which could be advantageous for raising knowledge of accessibility in engineering teams. However, one shortcoming of the method is that the list of barriers used in the evaluation was not validated with disabled users, which is a threat to the validity of the method.

2.3.3 User evaluation of web accessibility

Involving a representative set of users with disabilities in the evaluation of websites is a crucial need to perform effective accessibility evaluations. Even though many problems may be identified by means of inspection methods with experts, only tests with users are able to show the accessibility of a website. Evaluation with disabled users is considered the ultimate method for asserting the accessibility of websites (Petrie and Bevan 2009). However, it may not always be practical to evaluate all pages in a website with different types of users, tasks and environment conditions, especially given the difficulty of recruiting users with specific types of disabilities. Nevertheless, it is very important that the evaluation of crucial web pages in websites by disabled users be incorporated in the evaluation of websites, along with other expert-based evaluation methods.

Performing a user evaluation involves a number of steps (Monk 1993, Stanton 2005), including user recruitment, task preparation, running the evaluation and reporting the results.

The *user recruitment* task is a very important one to determine the success of a user evaluation. It is very common in many evaluation procedures that the recruitment process only targets users from a certain circle. It is very important that a representative sample of the actual users of a website is selected for the user experiment, with a wider range of experiences and different disabilities. The feedback from observing real users experiencing an application with their own assistive technology is fundamental for a true understanding of accessibility errors (Petrie and Bevan 2009).

Recruiting big samples of users with a varied range of disabilities may be a difficult task (Petrie et al. 2006). However, according to results obtained from a large study performed by the Disability Rights Commission (2004), there is a significant overlap between disability groups in terms of accessibility problems, which suggests that even having some key user groups can have a very positive impact in helping find problems.

Given the difficulty of recruiting disabled users, conducting remote user evaluations may be a viable option in some projects. A study conducted by Petrie et al. (2006) revealed that the quantitative data obtained in remote evaluations are comparable to those obtained in a laboratory. However, they also point out that the amount and richness of qualitative data are not likely to be comparable.

The *task preparation* is another important step for the user evaluation. It is important that a representative set of tasks be defined to cover the main aspects and features of the evaluated website. When writing the set of tasks, it is important to make sure that they will be understood by the target users. Tasks that take exaggeratedly long to be performed should be avoided.

When *running the evaluation*, it is important that special attention be given to ethical and practical aspects to collect important data. From the ethical point of view, it is important that the user be given all the important explanation about the procedures during the briefing, a proper consent form be given and a debriefing be performed by the end of the procedures. Regarding the procedures, a good testing protocol should be used to capture the important problems to be identified in the evaluation (Ericsson and Simon 1993). The use of the “thinking aloud” protocol is a good instrument to help know what the users are thinking whilst performing a task. Recording the magnitude of each problem may also be very helpful for the analysis.

After the evaluation is concluded, the *summarisation of results* is an important stage, which aims to provide a list of problems and their impacts from the users' perspective.

The following section presents a review of studies in the literature that report on evaluations performed by disabled users, describing current knowledge in the literature about the problems disabled people have on the Web. The section also presents studies that compared evaluation of websites by disabled users with technical web accessibility guidelines.

2.4 Research studies with evaluation of websites with disabled users

Despite the importance of having websites evaluated by disabled users, there are far fewer studies in the literature with user evaluations than technical evaluations of conformance to guidelines. This section presents some of the main studies involving evaluation of websites by disabled users encountered in the literature, presenting their main findings and methodological approaches.

The largest study with evaluation of websites by disabled users found in the literature was performed for the Disability Rights Commission (DRC) of Great Britain, in 2004 (Disability Rights Commission 2004). In this study, a panel of 50 participants with different disabilities evaluated a set of 100 websites. The user panel included participants with visual impairments (both totally blind and partially sighted), hearing impairments, motor impairments and specific learning difficulties, such as dyslexia. The study involved both laboratory tests and tests performed remotely by users. One of the main findings from the DRC study was that around 45% of the problems encountered by disabled users were not covered by the WCAG 1.0 guidelines. A more detailed discussion of the DRC study is presented in Section 2.4.3.

Most studies found in the literature focus on a single user group when performing evaluations of websites. By far, blind users have been the user group that has received the most attention in comparison to other user groups. Following, an analysis on web accessibility studies involving users with print disabilities is presented. Section 2.4.1 presents studies with visually-impaired users, Section 2.4.2 discusses studies with dyslexic users, and Section 2.5 presents studies that compared results from evaluation of websites by disabled users with technical web accessibility guidelines.

2.4.1 Web accessibility studies involving visually-impaired users

Studies have performed evaluations of websites by disabled users in order to define their own sets of accessibility guidelines.

The first large study involving evaluation of websites by disabled users encountered in the literature was performed by (Coyne and Nielsen 2001). Coyne and Nielsen derived a set of guidelines from a series of studies that involved 104 users, being 84 disabled users and 20 controls. The first part of the study was a qualitative study with 44 users (31 in the United States and 13 in Japan) that aimed to identify problems encountered by users on websites. The first study involved 35 visually impaired users and 9 users with motor impairments, evaluating 10 US websites and six Japanese websites.

The second part of the study performed by Coyne and Nielsen (2001) was a quantitative study that aimed to compare the performance of disabled users when compared to mainstream users. This second study included 20 blind users, 20 partially sighted users and a control group with 20 sighted users. Users had to perform four simple tasks (three on specific websites and a task using a search engine of their choice). The study analysed the success rate, time on task, number of errors and subjective rating. It was found that blind users could only succeed at 12.5% of the tasks, while partially sighted users succeeded at 21.4% and the control group at 78.2% of the tasks. The average time on task was 16min 46s for blind users, 15min 26s for partially sighted users and 7min14s for the control group. Partially sighted users had the highest average number or errors, with 4.5, followed by blind users with 2.0 and 0.6 from the control group. Blind users had a mean subjective rating 2.5 (in a 1-7 scale), while partially sighted users had 2.9 and the control group 4.6. This early study in 2001 showed that visually impaired users were very disadvantaged in comparison to sighted users.

Based on the two studies, Coyne and Nielsen (2001) derived a set of 75 design guidelines. The guidelines were grouped into the following groups: 1) graphics and multimedia, 2) pop-up windows, rollover text, new windows, and cascading menus, 3) links and buttons, 4) page organisation, 5) intervening pages, 6) forms and fields, 7) presenting text, 8) search, 9) shopping, 10) tables and frames and 11) trust, strategy and company name. The guidelines provide very good supporting evidence from the users' perspective on the types of problems encountered by users. However, they lack more detailed information regarding the frequency and severity of the problems encountered.

Theofanos and Redish (2003) performed an exploratory study with 16 blind participants using US government websites, using a think aloud protocol in a laboratory environment. Their study derived 32 guidelines from 16 facts observed in their study. Their findings supported design practices such as the use of a "skip to content" link in the beginning of the web page, but also showed that not all users will make use of it. The study grouped the findings based on observations about how users use their screen reader, how they navigate through websites and how they fill out forms. Theofanos and Redish's study is one of the earliest studies that provide website design recommendations for blind users based on an empirical study. However, the study does not make any comparisons between the data from their studies and their coverage by other sets of guidelines, such as WCAG 1.0.

Leporini and Paternò (2004) also proposed another set of 15 design recommendations for websites regarding blind users. The guidelines included issues such as not having too many links and frames, helping a user to identify a section in a page, identifying the importance level of different elements, questions related to the design of forms, among others.

A follow-up study was then conducted (Leporini and Paternò 2008) in order to test if the use of Leporini and Paternò's guidelines would improve blind users' performance when attempting tasks at websites. Two tests were performed with two different websites in each test, and two different groups of users. For each website, two versions were created, one that followed the proposed guidelines and one that did not. The first test was performed by 20 participants (10 blind and 10 partially sighted), and the second by 14 participants (14 blind and 6 partially sighted). The evaluations were performed remotely, using a tool to log users' actions and questionnaires to collect more information from the participants. The authors found that in both tests, the time to complete tasks was smaller in the website that followed their guidelines. Although the study presents some comments from participants about the usefulness of some of the guidelines, it is not possible to have specific information about the effectiveness of each guideline individually. A more detailed study with users in a laboratory setting could enable a more detailed analysis of the problems encountered in both websites, and how users interact with different components of websites in both cases.

Leuthold et al. (2008) proposed another set of guidelines to develop textual interfaces for blind users. They have a rather drastic approach to developing interfaces for blind users. They believe that the problems blind users have with interfaces are due to the use of graphical user interfaces (GUI) themselves. The authors in this paper propose a set of 9 guidelines to design separate enhanced textual interfaced tailored specifically to blind users. These guidelines tested in a study, consisting of an evaluation of 3 versions of a website: a) the original version, non-conformant to any guidelines, b) one textual version following their guidelines, c) and another version compliant to WCAG 1.0. The three versions were tested in an experiment by 39 blind users in a laboratory setting. No significant differences were found between the time to complete tasks, number of errors and user satisfaction between the original version of the website and the version that complied with WCAG 1.0. However, the authors found a significant difference in the time on task, number of errors and user satisfaction for search tasks between the original version and the textual version that complied to their guidelines.

The guidelines proposed by Leuthold et al. (2008) may be very useful for the design of websites for blind users. However, their proposal of building entirely separate websites for blind users may have some practical problems. In fact, Theofanos and Redish (2003) found in their study that many blind users encountered problems with websites with separate textual interfaces, as many companies who have designed a separate textual version of a website were not updated as frequently as the main graphical version. Nevertheless, the results encountered by Leuthold et al. (2008) suggest that offering effective personalisation features for blind users in websites can have positive effects in their interaction.

Mankoff et al. (2005) conducted a study comparing results of evaluations with blind users in a laboratory with the results of evaluations with automated evaluation tools, expert evaluations with and without screen readers, and remote usability evaluations by blind users. The baseline study consisted of the evaluation of 4 websites, with one task per website, by 5 blind users in a laboratory setting using the think aloud protocol. Participants in the laboratory study encountered 29 unique website problems in total. Following the baseline study, they performed another study with four different conditions. The same websites and tasks were evaluated by an automated evaluation tool, by web designers using WCAG 1.0 as reference (with and without the aid of a screen reader) and by a different group of blind users that performed the evaluation remotely. The web designers that took part in the study had little or no experience with accessibility, and they were divided randomly in the two groups (with and without screen reader). The panel of blind participants in the remote evaluation consisted of 9 experienced screen reader users. The results of the study showed that web designers using a screen reader were the group that found most of the problems previously identified in the baseline laboratory study. The authors report in the paper that they expected the remote evaluation to fare better than the other methods. However, they point out that there might have been some bias due to the different levels of expertise of blind users in the different conditions (laboratory and remote study).

In a very brief remark, Mankoff et al. (2005) also commented that there was no correlation between the severity levels assigned to problems by blind users in the laboratory study and the priority levels of related WCAG 1.0 checkpoints. This was an early finding about the problems with the priority levels that were later examined in more detail by other studies (Harrison and Petrie 2007, Petrie and Kheir 2007), in which it was confirmed that there were not strong correlations between the severity ratings of user problems and the priority levels of related WCAG 1.0 checkpoints.

Watanabe (2007) conducted a study on the impact of the use of headings on the navigation of websites by blind users. The study consisted of performing four tasks on two different versions of a website, being one version with headings properly marked up with HTML elements, and the other with headings just identified visually via CSS. Many blind users use special keyboard shortcuts to jump from heading to heading to have an overview of the structure of a web page, and this is only possible if the headings are properly marked-up. Watanabe's study involved 16 sighted and 4 blind participants. In order to counterbalance the order effect, half of the participants started with the marked-up version first, and half with the version without headings mark-up. The results showed that using proper heading mark-up reduced the disadvantage in the time taken to complete some of the tasks performed in the study when comparing blind and sighted users.

Babu and Singh (2009) performed a study with 6 blind users attempting to perform one task using a web-based learning environment, whilst "thinking aloud" as they performed their tasks. The authors coded each verbalisation from the users into single individual segments, which were then classified according to the stage they were in the task, following the Seven-Stages of Action model proposed by Norman (1988). The two main findings reported in Babu and Singh's study were problems related to the uncertainty about arriving on a new page and the susceptibility of skipping a question in the online assessment tool. The coding scheme based on the Seven-Stages of Action model was a very interesting approach used by the authors, particularly as the aim of the study was to understand the nature of the problems encountered by the user.

Unfortunately, the study was performed on a very small scale, and only two main problems were reported in the findings of the study.

Besides task-based evaluations with users, other studies have also performed surveys with disabled users to investigate what are the main problems that they encounter on websites. The first of these studies was performed by Lazar et al. (2007), in a survey with 100 blind users about what are the things that frustrate them the most when using websites. In this study, a time diary was used to record frustrating experiences that participants had when using websites at home. Every time they experienced frustration, they would fill out a questionnaire reporting on their experience. The study contained reports of 308 instances of frustration experiences. The top five causes of frustration reported in the study were: “a) page layout causing confusing screen reader feedback; b) conflict between the screen reader and application; b) poorly designed/unlabeled form; d) no alt text for pictures; and e) a three-way tie between misleading links, inaccessible PDF, and a screen reader crash”. Although the study collected information about the frustration experiences right after they happened, the time diaries cannot provide detailed information about the context in which the frustrating experiences took place in order to analyse it in detail.

The Web Accessibility In Mind initiative (WebAIM) conducted three online surveys with screen reader users (Web Accessibility in Mind 2009b, Web Accessibility in Mind 2009a, Web Accessibility in Mind 2011) to investigate their preferences and more information about their usage of screen readers. The first survey (Web Accessibility in Mind 2009b) had 1009 respondents, the second (Web Accessibility in Mind 2009a) 586 and the third (Web Accessibility in Mind 2011) 1107 participants. In all three studies, the results pointed that the most popular screen reader with blind users was Jaws. In the second study (Web Accessibility in Mind 2009a), respondents were also asked to point out what were the main accessibility problems they encounter in websites. The top five problems reported by the respondents in this study were: lack of "skip to main content" or "skip navigation" links (31.3% of participants), images with missing or improper descriptions (alt text) (15.9%), too many links or navigation items (9.6%), complex or difficult forms (7.1%) and missing or improper headings (6.6%). This showed the importance blind users place on problems related to navigation, description of images, difficult-to-use forms and proper use of headings.

Ruth-Janeck (2011a, 2011b) conducted an online survey with disabled users to investigate the main problems they encounter when using Web 2.0 applications, including media-rich applications and applications where users can collaborate with content, such as wiki systems. The study included 133 participants who were partially sighted, 124 blind, 96 with hearing impairments, 260 deaf, 75 with motor impairments and 89 with dyslexia. The problems were classified into four types of barriers: technical barriers, editorial and content-related barriers, designer barriers and organisational barriers. Technical barriers included issues such as having captchas, problems with error messages and forms in PDF. Editorial and content-related barriers included issues such as orientation and unclear arrangement of the page, bad descriptions of media content and bad names of links. Designer barriers included issues such as size of buttons and interactive elements and arrangement of links. Organisation barriers included issues such as problems with language support and quality of content.

The study conducted by Ruth-Janeck was performed with a very large number of users. A number of commonly encountered problems was reported in the study as well. However, many problem types have an unclear description. This could stem from the fact that the study was performed via a survey, and that there were no recordings of how users performed their tasks when they encounter such problems so they could be examined in more detail. Besides this, the explanation to the categorisation scheme adopted is not clear. The boundary between the different categories is unclear, and no evidence is given of the theoretical background that supports the categorisation scheme. A follow-up study was performed based on this investigation (Ruth-Janneck 2011b), with the comparison of the problems reported by the participants and the coverage of the problems by WCAG. This study pointed that most problems were covered by the guidelines. However, the findings could be questioned by methodological issues in the study design.

2.4.2 Web accessibility studies involving dyslexic users

In a recent literature survey of web accessibility and dyslexia, McCarthy & Swierenga (2010) reported that there is little work in the literature regarding the study of the accessibility of web sites for dyslexic users. The majority of the literature on dyslexia and web accessibility is related to guidelines to produce accessible web content to dyslexic users, derived from general guidelines for dyslexia.

A number of sets of guidelines have been produced to help developers produce more accessible web content for dyslexic users (Bradford 2005, British Dyslexia Association 2011, Kolatch 2000, Zarach 2002), as reported in a review undertaken by McCarthy & Swierenga (2010). Friedman and Bryen (2007) also conducted a review of 20 sets of guidelines from research studies and websites maintained by professionals and advocacy organisations connected to dyslexia and other cognitive disabilities, and compiled the guidelines most cited by these sources; most guidelines had to do with other cognitive disabilities, but some were applicable to dyslexia. Evett and Brown (2005) also performed an analysis comparing guidelines for producing accessible content for dyslexic and blind users, and reported to have found a high degree of overlap between guidelines for these two user groups.

With respect to empirical studies with dyslexic participants using websites, the largest study to date reported in the literature was conducted by the Disability Rights Commission of Great Britain in 2004 (Disability Rights Commission 2004). The study involved tests on 100 websites, performed by a panel of 50 participants, which included participants with dyslexia, visual, hearing and physical disabilities. Out of the 50 participants, 12 had dyslexia (Petrie et al. 2004). The study resulted in a total of 585 accessibility problems. In particular, the study found that the most recurring problems encountered by dyslexic users were: confusing page layout, unclear navigation, poor colour selections, graphics and text too small and complicated language.

Al-Wabil et al. (2006, 2007) conducted a study investigating navigation issues faced by dyslexic users. Their study comprised interviews with 10 participants with dyslexia. The participants were shown examples of web pages and asked to discuss about their experiences with navigation elements in web sites. Results pointed to how dyslexic users use search features, breadcrumb trails and other navigation resources. Although the study provided good insight from users' opinions, there was no empirical evidence from participants using real websites.

Rello et al. (2012) performed a study investigating layout preferences of dyslexic users using eye-tracking. The study involved 22 users and investigated eight aspects of text presentation: brightness levels in grey scale in writing, brightness levels in grey scale in background colours, colour contrast combinations, font size, character spacing, line spacing, paragraph spacing and column width. The study found that the influence of brightness in use of grey did not change how helpful users found the use of the colours. The preferred colour combinations from users were (background/foreground): yellow/black, white/blue, cream/black, white/black, yellow/blue and light mucky green/dark brown. The study also showed that most dyslexic users preferred larger font sizes than 12 or 14 pts. The study found that users preferred standard spacing between characters and larger spacing between paragraphs than between lines in paragraphs. Users also reported to prefer lines not to be long (60 to 70 characters) and to avoid narrow columns. Rello et al.'s study provides very interesting contributions based on empirical evidence with dyslexic users. However, the study was limited to layout-related issues and did not involve a complete task-based evaluation of websites.

Finally, Santana et al. (2012) performed a survey of common guidelines and techniques for dyslexic users encountered in the literature. They grouped the guidelines and techniques into nine groups: navigation, colours, text presentation, writing, layout, images and charts, end user customisation, mark-up and videos and audios.

2.4.3 The 2004 Disability Rights Commission Formal Investigation

The Disability Rights Commission (DRC) of Great Britain commissioned the first large-scale study that provided empirical data about the relationship between problems encountered by disabled users and technical web accessibility guidelines[13] in 2004 (2004). The result of this investigation was the largest known accessibility evaluation to date, with 1000 websites being evaluated with the accessibility module of WebXM™ developed by Watchfire[14] (IBM 2011) and 100 websites being chosen for expert and user evaluations. On the automated tests in the referred study, only 19% of the 1000 websites did not display any automatically identifiable violations of WCAG at Level A.

For the user evaluations, 913 tasks were undertaken over the 100 websites chosen, with participants being selected from a wide variety of people with disabilities. The user panel for the DRC study comprised of 50 users distributed in the following groups:

- blind people who use screen readers with synthetic speech or Braille output
- partially sighted people who may use screen magnification
- people who are profoundly deaf and hard of hearing
- people with specific learning difficulties such as dyslexia
- physically impaired people whose use of the Web may be affected by their lack of control of arms and hands, by tremor and by lack of dexterity in hands and fingers.

---

[13] The Disability Rights Commission of Great Britain is not extinct, and at the moment when this thesis was written, it had been aggregated to the Equality and Human Rights Commission.

[14] This tool was acquired by IBM and is now part of the IBM Rational Policy Tester Accessibility Edition

Each participant in the User Panel was asked to evaluate 10 websites, with two tasks per website. The study involved both tests in a usability laboratory and tests performed remotely. The tasks were distributed with 22% of the 913 tasks being performed in a usability laboratory, and 78% being performed remotely at home, with participants using their own equipment and software.

Regarding the completion of the tasks by the participants, the DRC study found a significant difference between different user groups. The users with the least percentage of task completion were blind users, who were only able to complete 53% of the tasks, followed by partially sighted users, who were able to complete 76%; dyslexic users completed 83% of the tasks, and physically and hearing impaired users, 85% of the tasks.

The difference between user groups in this study was also noted in the ratings of difficulty of the tasks. Each user was asked to rate how difficult he/she found each task, independent of whether the task was completed or not. The ratings were given in a scale of 1 = very difficult to 7 = very easy. Blind users were the ones who found the tasks “least easy” among the groups. The mean ease of task rating of blind users was 4.2, whilst partially sighted users had a mean rating of 5.1, followed by 5.6 by dyslexic users, 5.8 by hard of hearing impaired users, and 6.8 by motor impaired users.

A total of 585 accessibility and usability problems were identified in the DRC study, either by participants themselves in the remote evaluations, or by usability experts analysing video footage of laboratory tests. The main problems encountered by each user group in the DRC study are listed in the following paragraphs.

One of the important features of the results of this study is that categories of problems occur in more than one user group. The following is a list of problems encountered with their associated user groups:

- Confusing and disorienting navigation mechanisms (all groups)
- Unclear and confusing layouts of pages (all groups except blind users)
- Inappropriate use of colours and poor contrast between content and background (all groups except blind users)
- Graphics and text size too small (all groups except blind users)

There were a few problem categories that were unique to user groups:

- Blind users:
  - Incompatibility with screen reading assistive technology
  - Incorrect or non-existent labelling of links, form elements and frames
  - Cluttered and complex page structures
  - ALT tags on images non-existent or unhelpful
- Partially Sighted Users
  - Incompatibility with screen magnification software
- Deaf and hard of hearing users
  - Lack of alternatives for sound based media
- Dyslexic users
  - Complicated language or terminology

Given the types of problems encountered by different user groups, if it is not possible to have a panel with users with different types of disabilities, including users from some users groups will potentially cover many problems that are shared by other groups as well. It should be noted, however, the importance of including blind users in the evaluation of accessibility, as this user group has the majority of problems that are particular to them.

2.5 Studies on the relationship between evaluation of websites by disabled users and technical web accessibility guidelines

There is little investigation on the relationship between existing web accessibility guidelines, particularly the mostly widely used guidelines from the W3C, and results from evaluation of websites by disabled users. The small-scale studies performed by Leuthold et al. (2008), Leporini and Paternò (2008) and Theofanos and Redish (2003) derived their own sets of accessibility guidelines based on usability studies with blind users. However, those studies did not present any comparison between the results from user evaluation of websites and the existing guidelines proposed by the W3C's WAI.

Other studies provided empirical evidence on the relationship between the WAI's web accessibility guidelines and the evaluation of websites by disabled users. These studies included analyses on whether the problems encountered by disabled users had related guidelines in WCAG and whether there was or not any relationship between the problems' severity levels assigned by users and the priority level of related guidelines. These studies provided very interesting insights, and also raised important research questions that helped motivate the work reported in this thesis. Following, the main contributions and limitations of these related studies are described in details.

2.5.1 DRC

Concerning the relationship between WCAG 1.0 and the problems experienced by the users, the DRC study (Disability Rights Commission 2004) found that only 55% of the problems were related to a checkpoint in WCAG 1.0. The study points out that 45% of the problems could be present in any WCAG 1.0 conformant website, irrespective of conformance level.

The study also found that, from the 55% of problems that were related to a checkpoint in WCAG 1.0, 82% of them were related to a set of only 8 of the 65 checkpoints in WCAG 1.0. However, only 3 of those checkpoints were assigned priority 1 in WCAG 1.0. The problems related to the other 5 checkpoints accounted for 63% of user problems related to a WCAG 1.0 checkpoint, but they could potentially be present in any website conformant to WCAG 1.0 just at level A, since they had priority levels 2 or 3.

In a response by the WAI to the claims in the DRC study (Brewer 2004), the WAI argued that many of the problems told to be not covered by the WCAG 1.0 would be covered by the User Agent Accessibility Guidelines 1.0 (UUAG). More detailed information about the circumstances when the problems occurred would have allowed for more clarification about whether the problems were related to issues related to the implementation of the websites or with the assistive technology being used.

Another point argued by the WAI's response was the lack of a representative number of WCAG 1.0 conformant websites in the sample. More detailed analyses could be performed in an evaluation of websites with more variability in the levels of conformance to WCAG guidelines, as well as variability in the number of different checkpoints violated and instances of violations.

It is important that in light of these criticisms that in order to detail the nature of problems encountered by disabled users the studies in this thesis must:

- perform evaluations with disabled users in laboratory with recordings of the sessions, in order to enable more detailed analyses of the types of problems encountered;
- archive pages visited by users during the evaluations, in order to allow analysis of the source code of web pages where problems were encountered by users to determine the technical causes of problems;
- have more variability in the sample of websites evaluated in terms of conformance to WCAG, and numbers of different checkpoints violated and instances of violations.

The conformance analyses performed in the DRC study also provided other important findings that were used in the method of the present work. Audits of the websites revealed a very high correlation (r > 0.9) between the number of violations of WCAG 1.0 checkpoints in the home pages of websites and the number of violations in inner pages of websites (Petrie, H. Personal communication, 26/09/2012)[15]. This was applicable both for the measures of how many *instances* of violations of checkpoints occurred in web pages, and the number of *different checkpoints violated.* This finding was important to support the method to sample websites to be used in the present study, based on the audit of the home pages of websites.

2.5.2 Study on severity ratings by Harrison and Petrie

Harrison and Petrie (2007) conducted a study to analyse the correlation between the severity ratings assigned by users and usability experts, and the priority levels assigned to related guidelines. The study comprised the evaluation of 6 websites, being 3 commercial and 3 governmental websites. The websites were evaluated by 6 participants, being 2 visually impaired, 2 dyslexic and 2 non-disabled users as controls. Each participant attempted 2 tasks on each website.

---

[15] This information was obtained from Prof. Helen Petrie, based on the results in detailed reports from the DRC study that contained further information not published in the main report of the study in - Disability Rights Commission (2004) *The Web: access and inclusion for disabled people: A formal Investigation conducted by the Disability Rights Commission*, London: The Stationery Office.

The study yielded a total of 71 problems experienced by the users. Each time a problem was encountered, the user was asked to rate the problem in a scale ranging from cosmetic, minor, major or catastrophic. A usability/accessibility expert later rated the same problems independently.

The study found that there was a significant correlation between the severity rating assigned by users and the severity rating assigned by the usability/accessibility expert, with agreement in 69% of occasions. It was also found, though, that in average, users tended to rate problems less severely than experts.

It was also found in this study that only 22 of the 71 problems experienced by users were related to a WCAG 1.0 checkpoint. No correlation was found between the priority levels of related checkpoints in WCAG 1.0 and the severity rated assigned by either the experts or the users. For example, users identified 6 accessibility problems rated as catastrophic problems, of which 4 were related to priority 2 checkpoints, and 2 were related to priority 3 checkpoints. It would be expected that problems that are rated so severely by users would be assigned the highest priority in the accessibility guidelines.

The study performed by Harrison and Petrie provided some important insight into the research question regarding the relationship between technical web accessibility guidelines and evaluation by disabled users. It raised the issue that the priority levels in the guidelines, a cornerstone in the development of policy and prioritisation of guidelines, does not have any correlation with the severity ratings assigned by users or accessibility experts.

As the authors point out in their paper, larger studies with a wider variety of websites and users would be necessary to confirm the findings and to perform more detailed analysis on the priority levels and severity ratings from users and accessibility experts.

2.5.3 The relationship between usability and accessibility of websites by Petrie and Kheir

A study conducted by Petrie and Kheir (2007) investigated the relationship between usability and accessibility of websites by conducting usability and accessibility evaluations of websites by disabled and non-disabled users. The study also presented interesting findings about the relationship between severity ratings of problems encountered by disabled and non-disabled users and the priority levels assigned by usability and accessibility guidelines.

Petrie and Kheir's study consisted of an evaluation of the websites of two mobile phone companies. The study included a panel of 6 blind participants and 6 sighted participants. Participants were matched as far as possible on age, gender and general computer and Internet experience and expertise. The websites evaluated were Orange (www.orange.co.uk) and T-Mobile (www.tmobile.co.uk) in 2006, with 7 equivalent tasks being attempted by each participant on each website.

For the Orange website, their study yielded 168 instances of problems encountered by blind users, and 90 problems encountered by sighted users. Besides encountering more problems than sighted users, blind users also had a significantly lower success rate in completing their tasks (50.7% for blind users versus 70.2% for sighted users). For the T-Mobile website, blind participants encountered a total of 120 instances of problems, whilst sighted participants encountered 102 instances of problems. On the T-Mobile website, there was also a significant difference in the success rate on tasks, with 66% of tasks completed by blind participants, versus 83% completed by sighted users.

An analysis on the severity ratings of the problems encountered in the two websites was also performed in their study, comparing the severity ratings assigned by participants to those assigned by researchers and by accessibility and usability guidelines. The study included comparisons with the WCAG 1.0 priority levels and the importance level assigned by the Research-based Web design and usability guidelines from the Department of Health and Human Services (HHS) (Koyani et al. 2004). When a problem was found by more than one participant, the mean severity rating was taken as a measure of the participants' rating. In a similar manner, when a problem had more than one relevant guideline associated with it, the mean priority level/importance level of the guidelines related to a problem was taken as the measure for each of the two sets of guidelines (WCAG and HHS). A researcher also independently assigned a severity level to each problem encountered, without access to the participants' ratings.

Their study found that there was a small correlation between the severity ratings assigned by participants and the ratings assigned by researchers. They found a significant correlation between the severity ratings assigned by blind participants and the importance level assigned by the HHS guidelines on both websites, but the correlations were in the opposite directions to the predicted, which means that participants tended to give higher ratings for problems with lower ratings on the HHS guidelines, and vice-versa (Petrie and Kheir 2007). Regarding the WCAG guidelines, the study did not find any significant correlation between ratings from blind participants and the priority levels in either of the websites tested. These findings pointed out to very worrying concerns about how valid the priority and importance levels in guidelines are valid, and show that considerably more work needs to be done to establish valid priority levels based on how problems impact users.

The main contribution of Petrie and Kheir's paper, though, was on the analysis between the usability and accessibility of the websites, and exploring the relationship between the two concepts. The study found that the problems encountered by sighted and blind users were overlapping sets. On the Orange website, 66% of the problems were encountered only by blind participants, 17% by sighted participants only, and 17% by both blind and sighted participants. On the T-Mobile website, 57.5% of the problems were encountered only by blind participants, 31.9% only by sighted participants, and 10.6% by both blind and sighted participants.

The analysis of the severity ratings of problems encountered on the T-Mobile website found that there was a significant difference between the ratings from blind and sighted participants. On this website, problems that were encountered by both groups were rated significantly more severely by blind participants than by their sighted counterparts. This finding suggests that evaluation with disabled users can help identify problems that can also be found by non-disabled users, which will be "amplified", in a sense that they can affect disabled users more severely.

Regarding the definition of accessibility, the paper's findings revealed that accessibility problems were not a complete sub-set of usability problems, as suggested by Thatcher et al (2003), and that usability problems were not a complete sub-set of accessibility problems, as might be inferred from Shneiderman (2000, 2003).

The contributions from Petrie and Kheir's (2007) study were very significant, and many aspects of the methodology applied in their study were used to orient the method used in the study reported in this thesis. The conclusions about the relationship between usability and accessibility are also very important, as they provide valuable empirical data to provide a better understanding between the boundaries between usability and accessibility.

2.5.4 Comparison between user evaluation and WCAG 1.0 and WCAG 2.0 by Rømen and Svanæs

Rømen & Svanæs conducted two studies (Rømen and Svanæs 2008, Rømen and Svanæs 2011) in which they aimed at validating whether problems encountered by disabled users were covered by WCAG guidelines. These studies are very closely related to the study presented in this thesis. Following, the main findings of the two studies are discussed, as well as some shortcomings from the studies that were addressed by the work reported in this thesis. The first study (Rømen and Svanæs 2008) was published before the publication of the WCAG 2.0 (Caldwell et al. 2008), and hence only investigated the coverage of user problems by WCAG 1.0. The second study expanded on the analysis of the evaluations to include coverage by WCAG 2.0.

Both studies were based on the evaluation of 2 websites of 2 municipalities in central Norway. Four equivalent tasks were defined for each of the two websites, most related to basic tasks citizens would undertake in a municipality governmental website.

Their study included 7 disabled participants and 6 non-disabled participants. The group of disabled participants included 2 totally blind, 1 severely visually-impaired, 2 motor-impaired and 2 dyslexic participants.

A total of 176 instances of problems were identified by the participants. These problems were related to a total of 80 website problems, that may have occurred to different participants. When comparing problems identified by each group, the study found that 62% of the problems were encountered only by disabled users, 25% only by non-disabled users, and 14% by both groups.

An analysis on the correlation between the severity of the problems and WCAG priority levels was performed in the study as well. The severity ratings were assigned based on Molich's (2007) rating scale, ranging between *cosmetic* (makes it slightly harder for user to complete task), *serious* (considerably slows user down) and *critical* (prevents user from completing the task). It is not clear from either of the papers describing the study whether the severity ratings were assigned by the participants or by the researchers. However, both papers (Rømen and Svanæs 2008, Rømen and Svanæs 2011) mention that the severity ratings were assigned to *website problems*, and not to *problem instances* as they were encountered by users. If ratings had been assigned by users, they would have been assigned in each occurrence of a problem instance. This seems to suggest that the severity ratings discussed in the papers were assigned by the researchers *a posteriori* during the analysis of the sessions. The 47 problems that were only encountered by disabled users were classified into 6 critical, 18 serious and 23 cosmetic problems.

The study also tried to match each of the website problems with a WCAG 1.0 checkpoint and WCAG 2.0 success criterion that could have identified the problem in an expert evaluation. It was found that only 27% of the problems could have been identified by WCAG 1.0 checkpoints, and 35% by WCAG 2.0 success criteria. The authors also found that a combination of WCAG 1.0 and WCAG 2.0 would make for a coverage of 38% of the website problems encountered by users.

Neither of the papers on this study reported any numerical correlation index between the priority levels and severity ratings. However, the authors argued that they did not find any correlation between severity ratings and priorities, by showing that, for example, out of 6 problems rated as *critical*, only 1 problem had a related WCAG 1.0 checkpoint at priority 1, and only 2 at level A in WCAG 2.0.

The authors concluded their study stating that there was a slight improvement from the coverage of problems from WCAG 1.0 to WCAG 2.0, but that the percentage of problems covered by the guidelines was still very low.

Rømen & Svanæs' study provides very interesting insights, and confirms problems with the coverage of WCAG guidelines identified in previous studies (Disability Rights Commission 2004). It also confirms the existence of problems with the lack of correlation between severity ratings of user problems and the priority levels assigned by guidelines (Harrison and Petrie 2007, Petrie and Kheir 2007).

However, as it was also pointed out by the authors (Rømen and Svanæs 2011), the study still had some limitations that could not provide more details about the relationship between problems encountered by users and WCAG guidelines. The authors recognise that the study was performed with a small sample of users. Besides the sample of users, the two websites evaluated were very similar, and there was not enough variability in the technologies used in the websites and in their levels of conformance to WCAG. Having a more varied sample of websites to be tested could provide more detail about the coverage by the guidelines of different technology such as multimedia content, interactive applications, different navigation structures, and others. Furthermore, having a sample with websites at different conformance levels with WCAG could enable analyses into whether conformance to the different levels of WCAG can lead to any impact on the problems that disabled users experience on websites. Regarding the method, it was a shame that the papers suggest that severity ratings were not obtained from users. It would have been very interesting to be able to perform analyses on user severity ratings, in the line of the studies performed by Harrison and Petrie (2007) and Petrie and Kheir (2007).

2.5.5 Summary of studies on the relationship between evaluation of website by disabled users and technical web accessibility guidelines

Section 2.5 presented the main studies that compared the evaluation of websites with users and with technical web accessibility guidelines. Most studies pointed out problems with a number of user problems not being covered by the main accessibility guidelines defined by WCAG 1.0 and WCAG 2.0 (Disability Rights Commission 2004, Harrison and Petrie 2007, Petrie and Kheir 2007, Rømen and Svanæs 2008, Rømen and Svanæs 2011). They also showed problems with the lack of correlation between the severity of problems assigned by users and the priority levels assigned to problems by the guidelines. One study suggested that most problems were covered by WCAG 2.0 (Ruth-Janneck 2011a, Ruth-Janneck 2011b). However, there are serious validity problems with the analysis, since the study did not map user problems with WCAG 2.0 success criteria, which are the testable statements, but with broader and more vague guidelines.

Although there is evidence showing problems with current accessibility guidelines, the studies presented still had some limitations that prevented them from answering to more specific research questions about the relationship between user problems and technical web accessibility guidelines. The study presented by Ruth-Janeck (2011a, 2011b) was not conducted with a task-based approach, but by asking participants about problems they commonly have. Other studies had some limitations regarding the samples of websites and small sample of participants (Harrison and Petrie 2007, Petrie and Kheir 2007, Rømen and Svanæs 2008, Rømen and Svanæs 2011). The main study in the area, conducted by the Disability Rights Commission of Great Britain (2004), had a very large sample of users and of websites. However, there were still some limitations with the study. One of the limitations was the lack of variability of the conformance levels of the websites selected for the evaluation, as there were very few websites conformant to even the lowest level of WCAG 1.0. The users in the study performed a large amount of the tasks in the study by means of remote evaluation. Although the evaluation provided a substantial body of quantitative data about the problems encountered, some issues could not be identified due to the lack of more details. For example, one of the main problems encountered by blind users was classified as “incompatibility between screen reading software and web pages”. In a study performed in a laboratory with video recording, it would have been possible to examine into further detail to understand the nature of these problems.

The study presented in this thesis built up on the previous work presented in this section, and included new elements to overcome some of the limitations of the related works presented.

2.6 Summary of the chapter

This chapter presented the main concepts and a review work related to web accessibility and its evaluation. The chapter presented studies that have investigated the accessibility of websites by involving disabled users in the evaluation, and other studies that only performed evaluations based on technical web accessibility guidelines.

A review of related studies that performed evaluation of websites by disabled users was performed, along with studies that compared evaluation of websites by disabled users and technical web accessibility guidelines was also performed. This review presented the main concerns raised by these studies in relation to the lack of empirical evidence supporting current technical web accessibility guidelines, as well as indications of problems with the coverage of problems encountered by users by the guidelines. Limitations in the studies reviewed were also discussed.

The literature review presented in this chapter presents the context in which the work presented in this thesis is inserted. The limited number of studies of the accessibility of websites involving disabled users and the questions raised by related studies on the relationship between user evaluation of web accessibility and technical guidelines were important motivations for the development of the present work. Limitations in the methodology of previous work presented in this chapter were considered in the development of the method for the development of the work presented in this thesis.

Chapter 3. Method

This chapter presents a description of the methods used to conduct the study presented in this thesis, in order to address the research questions presented in Chapter 1. The research questions set for this study require both methods involving the evaluation of websites by disabled users and audits of websites by accessibility experts using technical guidelines. Methods for both of these aspects are presented in this chapter.

3.1 Design

The main research study presented in this thesis was an empirical study conducted with users who have different print disabilities, specifically blind, partially sighted and dyslexic who undertook tasks on a variety of websites. The main objective of this study was to collect a corpus of problems encountered by users with print disabilities that could be analysed regarding the type and frequency of problems, compared between user groups and examined in terms of how the problems are addressed by existing web accessibility guidelines.

Instead of conducting several small studies, it was decided to perform a single larger study with the same set of websites for different user groups. This would allow for comparisons between the groups of the characteristics of problems they encounter, the frequency of those problems and the perceived severity to the user groups.

It would be ideal to include as wide a range of disabilities as possible, including participants with visual, hearing, physical disabilities and specific learning difficulties, such as dyslexia. However, due to time and resources limitations, it would not be possible to include a representative number of participants in all those user groups. The study included users who were totally blind, users who were partially sighted (had some vision impairment, but who still had some vision), and users with different types of dyslexia. Those users have different ways of interacting with websites and may encounter different types of problems. Blind users use screen reader software that synthesises textual content on the screen in the form of speech, and normally use only a keyboard to perform data input. Partially sighted users can use a range of different technologies and adaptations, such as screen magnification software to change the size and/or colour scheme, or can use special settings on their operating system or web browser to perform these changes. Some partially sighted users also use speech synthesis to help them read content on the screen. Dyslexic users can encounter problems related to reading/decoding text. Some users may need to change settings related to the presentation of text, such as size, colour, spacing or alignment, or use speech synthesis software to help them read text as well.

This study did not include non-disabled users as a control group. The main aim of this study was not to differentiate usability problems encountered by disabled and non-disabled users, as performed by previous studies (Petrie and Kheir 2007). The study focused on disabled users specifically, and investigating the kinds of problems that they encounter on websites.

The independent variable used in the study was the conformance of websites to WCAG 1.0 and WCAG 2.0. As previous studies have demonstrated a high correlation between the violations of accessibility guidelines in the home page and the violations of accessibility guidelines in the other pages in a website, as reported in Section 2.5.1 in the description of the DRC study (2004), a sample of websites was created from across these different conformance levels as determined by a conformance audit of the home page of the website. The websites were selected from domains such as government, education entertainment and commerce. The websites were also selected to ensure that they had different types of interactive components.

Although WCAG 2.0 had been published in 2008, at the time when this study design was laid-out (in 2009), WCAG 1.0 was still in use by many organisations and in governmental legislation. For Australian governmental websites, for example, WCAG 2.0 became mandatory only in July 2010 (Australian Government 2010), and only from January 2010 in websites of the European Commission (European Commission 2012). In Brazil, a new set of web accessibility guidelines that incorporated WCAG 2.0 was only published in 2011(Brazilian Government 2011). For this reason, the analysis in the present study still considered conformance to WCAG 1.0. This also aimed to enable comparisons of the relationship between problems encountered by disabled users and the two versions of the guidelines, and comparisons with previous related studies that used WCAG 1.0, such as the DRC study (Disability Rights Commission 2004).

A set of tasks was identified for each website for users to undertake during the study. These tasks were naturalistic, in that they were typical things that users would do when visiting websites in their own time.

Users undertook tasks on up to 10 different websites while being observed in a laboratory setting. Given time that it takes to perform the tasks on websites, especially for blind and partially sighted users, it was not possible to have the same users evaluate all websites in the selection. Besides the issue with time, different participants were able to evaluate a different number of websites during the time they had available. Due to those reasons, websites were evaluated in cycles for each disability group, and the order was reshuffled at each cycle to avoid any ordering or fatigue effects. Table 3.1 presents an example of how the order of the evaluation would be with four websites.

Table 3.1. Example of cycles of evaluation of websites by users to avoid order effect

Cycle	First website	Second Website	Third Website	Fourth website
1	A	B	C	D
2	C	A	D	B
3	D	C	B	A
4	B	D	A	C

During the tasks users applied a concurrent "think aloud" verbal protocol (Ericsson and Simon 1993). Following this protocol, participants were asked to verbally express what they were thinking as they approached the tasks they were attempting to accomplish on the websites. The participants' comments about what they were thinking would provide more insights about the users' mental model of the websites and their plans of action when trying to perform the tasks. When users encountered a problem, they stopped and described the problem to the evaluator in their own words. They also provided a rating of the severity of the problem in terms of how it would affect their completion of the task. The ratings available to the users were as follows:

1. Cosmetic: an irritating problem that they overcome easily
2. Minor: a problem which will stop the user for a short period of time or will be overcome relatively easily
3. Major: a problem which will stop the user for a long period of time or will be difficult to overcome and continue the task

4. Catastrophic: a problem which stops the user from continuing

These problems and their ratings were the major dependent variables collected during the study.

Besides the description of the accessibility problems encountered by users and their severity ratings, other variables related to the performance and satisfaction with the websites were also analysed. The main variables to be analysed regarding the tasks undertaken by users were:

- Problems encountered by users and their severity
- Task completion rates
- Difficulty to perform each task

Those measures provided a detailed picture of the problems print-disabled users encountered and how they affect their usage of websites. Along with information about problems and their severity, task completion rates, ratings of difficulty to perform tasks, common measures used in usability studies, provide important information about how accessibility problems affect disabled users in their tasks on websites.

The chosen means to measure the difficulty to perform each task was to ask participants to rate it in a 5-point Likert-scale, where 1 means “very easy” and 5 means “very difficult”.

The corpus of problems for each user group was analysed to categorise the problems. The final categories of problems were analysed for their frequency of occurrence and overall severity for the user groups. These problems were then compared to WCAG 1.0 and 2.0 specifically looking at if there were specific guidelines that addressed the problems. If guidelines were identified that addressed the problems, then it was also investigate if the pages on which the problems were encountered implemented any of the techniques recommended to address them.

3.2 Accessibility audit processes

The careful selection of websites for this study was fundamental to enable a good analysis of accessibility problems experienced by users on websites. These websites had to be from different contexts with different technologies and resources commonly used on websites. It was also very important that the selected websites had a good enough variability in terms of their conformance to web accessibility guidelines.

In order to achieve this goal in the selection of websites, a careful selection process was performed, by means of audits of the home page of hundreds of candidate websites. The following sections describe the procedures used to perform accessibility audits for this selection – for both WCAG 1.0 and WCAG 2.0 – and the final sample selected for the study.

3.2.1 Automated and manual inspection tools

The automated tool selected to help the audits was the tool Hera (Benavídez et al. 2006). This tool was selected because it was available for free use online, and because it presents useful features to highlight particular elements on a web page that should be analysed manually for particular CPs, as well as presenting a well organised report for the automated evaluations.

Besides Hera, other tools were also used for specific CPs. The Accessibility Evaluation toolbar for Firefox (Pederick 2011) was used for several tests, including the verification of text alternatives for images, highlighting tables, forms, headings and other elements. Figure 3.1 shows a screenshot of the web developer tool bar, with an example of a page with its headings highlighted by the tool. Pages were also tested with the screen reader Jaws, in order to verify issues that blind users encounter. In particular, the tests with Jaws aimed to verify the order in which screen readers would read pages, using the “links list” to see if links made sense when listed out of context and if pages contained headings.

Figure 3.1 – Screenshot of Firefox web developer tool bar, with example of feature to highlight headings

The verification of colour contrast was performed with the aid of the tool Juicy Studio Colour Contrast Analyser (Juicy Studio 2012). The tests for colour contrast were performed using the algorithm to test luminosity levels and colour difference, as described in the "Techniques For Accessibility Evaluation And Repair Tools" (Ridpath and Chisholm 2000). Although this algorithm never became a W3C recommendation, it was the "de facto" standard used to test colour contrast with WCAG 1.0 until the publication of the new algorithms used in WCAG 2.0. The HTML (Hypertext Mark-up Language) and CSS (Cascading Style Sheets) conformance verification was performed using the W3C HTML validation service (World Wide Web Consortium 2012) and W3C CSS validation service (Hégaret and Smeman 2012).

3.2.2 Procedure for web page audits with WCAG 1.0

This section describes the procedure for the accessibility audits of web pages using the Web Content Accessibility Guidelines (WCAG) 1.0 (Chisholm et al. 1999).

A protocol was developed to perform the audits with WCAG 1.0 in the context of this research based on the conformance evaluation process defined by WCAG 1.0[16] and including specific procedures to test each checkpoint. The WCAG 1.0 documents do not contain any explicit definition of a set of tests to attest conformance to the CPs. The tests included in this protocol aimed to cover as much as possible of the requirements described in each of the WCAG 1.0 CPs. The test procedures were drawn from techniques described in the W3C document "Techniques For Accessibility Evaluation And Repair Tools" (Ridpath and Chisholm 2000) and some tests developed in the context of the BenToWeb Project (Benchmarking Tools and Methods for the Web), (Velleman et al. 2007) part of the Web Accessibility Benchmarking Cluster of European Projects.

For each CP, one or more tests were performed to determine whether a web page was conformant or not, depending on the requirements described for the CP. A web page was given a "pass" on a CP if it passed all the tests performed for that particular CP.

During the execution of each test, the number of instances of violations of each CP was also recorded, counted as the number of instances in which each test applied to a specific CP failed. For most tests, the definition of an instance of a violation was

[16] Available online at http://www.w3.org/WAI/eval/conformance.html, last accessed 24/09/2012

applied as the number of interface components that failed to comply with a test for a CP. In the case of CP 3.2 – "Create documents that validate to published formal grammars", WCAG 1.0 is not clear about how to count the number of instances of violations. The number of instances can be counted either as each individual violation of the HTML specification or a general fail/pass depending on whether a given page passed a validation test. The tests for CP 3.2 defined in the BenToWeb Project[17] are defined to count the HTML validation test as a single violation. This was the definition of the number of violations of CP 3.2 in this study, as it is the closest to the definition of the checkpoint in WCAG 1.0.

3.2.2.1 Limitations and Inconsistencies in WCAG 1.0 checkpoints

In cases where WCAG 1.0 did not define a clear test procedure for specific checkpoints, it was necessary to resort to other procedures commonly used in practice and developed by other sources.

For CP 2.2 (colour contrast), the WCAG 1.0 techniques do not provide any recommendation about the level of contrast required. The audit protocol included a test with the commonly used luminosity and colour difference test available at the set of tests for automatic evaluation tools (Ridpath and Chisholm 2000).

For CP 14.1 ("Use the clearest and simplest language appropriate for a site's content"), there was no indication of how to test how easy it was to read a document. In this case, the audit protocol used the more detailed procedure developed in the context of the BenToWeb Project (Velleman et al. 2007), that involved several detailed checks on issues related to the readability of texts in English.

Commonly used technologies, such as Flash and PDF are not allowed in WCAG 1.0. In order to ensure that precise comparisons between WCAG 1.0 and WCAG 2.0 could be performed, in the audit protocol with WCAG 1.0, the use of such technologies without alternatives as recommended by WCAG 1.0 would still be counted as violations, in keeping with the wording of the old version of WCAG 1.0.

[17] Available online at http://www.bentoweb.org , last accessed 24/09/2012

3.2.3 Procedure for web page audits with WCAG 2.0

This section describes the procedure for the accessibility audits of web pages using the Web Content Accessibility Guidelines (WCAG) 2.0 (Caldwell et al. 2008), adapting the same conformance evaluation process used for WCAG 1.0 with specific particularities from changes in the guidelines.

When this audit protocol was developed, there were still few evaluation tools available and there still seemed to be some problems related to the understanding of the rules for conformance in the new recommendations and how they are implemented during accessibility audits by many evaluators (Alonso et al. 2010, Brajnik et al. 2010).

As discussed in Section 2.2, WCAG 2.0 has a different structure from that of WCAG 1.0. The set of 61 WCAG 2.0 guidelines are organised into four principles, that indicate that content should be *perceivable*, *operable*, *understandable* and *robust*. Each guideline has a number of *success criteria* (SC), which are the statements that actually guide the auditing process with WCAG 2.0.

Following the rules for satisfying SC in WCAG 2.0, as described in Section 2.2, the method for audit with WCAG 2.0 in this study had the sufficient techniques as the starting point to evaluate conformance to WCAG 2.0. For each SC, the test procedures listed with the SC were applied as a first step in the audit.

The number of different techniques, conditions in which they are applied and logical relationships between them make it very difficult for evaluators to understand it, particularly as some parts of those rules are contained in separate documents. In order to make it easier to perform the audits with WCAG 2.0, a checklist was developed containing an overall view of all those rules in one single document.

In case a web page failed to pass the set of sufficient techniques for a given SC defined in WCAG 2.0, it was verified if there was any evidence of implementation of accessibility features that would satisfy the SC other than those suggested in the list of sufficient techniques. However, in the absolute majority of cases, web pages that failed sufficient techniques also failed to have other provisions that would meet WCAG 2.0 SCs. Exceptions for this were in cases when technologies such as PDF (Portable Document Format) or embedded videos were used. At the time the audits were performed, there were not any sufficient techniques for PDFs, for example (Cooper et al. 2010b). Techniques for PDF documents were only made available in January 2012 (Web Accessibility Initiative 2012). In the particular case of audits of PDF documents, it was checked if the document was properly tagged and if appropriate implementations for the requirements in the SCs were made, such as providing alternatives to images.

In the same manner as in the audits with WCAG 1.0, during the execution of tests for each technique, the number of instances of violations of each SC was also recorded, counted as the number of instances in which each test applied to a specific SC failed. This was only counted for techniques that were applicable to each specific element in the page. In the case where a given component failed one sufficient technique, but passed another sufficient technique or set of sufficient techniques that would be enough to pass the SC, the failures to the technique not successfully implemented would not be counted. For example, for SC 2.4.5, suppose a website had a search feature (technique G161), a table of contents (technique G64), but did not have a site map (technique G63). In this example, the website would have passed SC 2.4.5, as only two of the recommended sufficient techniques are enough to meet the SC. In this case, the failure to provide a site map would not be counted as a violation.

3.2.3.1 Limitations and inconsistencies in WCAG 2.0

Following, the main inconsistencies and difficulties encountered with the interpretation of techniques and their arrangement for sufficiency, and how these inconsistencies were resolved for the audits with WCAG 2.0 performed are presented.

Following the test procedures of some techniques in WCAG 2.0 can be a problem for evaluators when the technique does not contain all the details about when it should be tested and how to establish whether each test passed or not. In many techniques, the details about which elements to test and key elements for the test procedures are contained in sections external to the techniques. In such cases, the audit protocol aggregated all important information in one place.

One example of this problem is with technique G145 – “Ensuring that a contrast ratio of at least 3:1 exists between text (and images of text)”. The test procedure for this technique contains the formulae for measuring “relative luminance”, and the final outcome of test is stated as “Check that the contrast ratio is equal to or greater than 3:1” (Cooper et al. 2010b), but did not include constraints about the text size where it is applicable, that was in a description external to the technique.

Another very frequent problem was related to the same technique being used in very different contexts, with different test procedures for each of them. WCAG 2.0 attempted to give developers the opportunity to choose the degree of thoroughness in the implementation of some accessibility features, depending on which conformance level they want to achieve. For example, SC 1.4.3 (level AA) and SC 1.4.6 (level AAA) are both concerned with the colour contrast of text or images of text. However, the requirements for meeting 1.4.6 are tighter than the requirements for meeting 1.4.3.

The approach of having different SCs related to the same issue was also used when exceptions were made for success criteria at lower conformance levels. SC 1.4.9 (level AAA) states that "images of text are only used for pure decoration or where a particular presentation of text is essential to the information being conveyed", except for essential images, such as logotypes. Related SC 1.4.5 (level AA), though, makes a further exception for images that can be customisable. The problem with those two success criteria arises from the fact that, despite having different exceptions, the techniques and test procedures for both are exactly the same. In such cases, the restrictions that were not in the technique were specified in the audit protocol to ensure they were addressed.

3.3 Materials

3.3.1 Websites and tasks

The home pages of a subset of 72 live websites of the 100 websites used in the original DRC study were audited for their accessibility. Only the home page was audited as previous studies (Disability Rights Commission 2004) have established a very high correlation of the WCAG 1.0 conformance of the home page of a website with the WCAG 1.0 conformance of other pages of that same website, as described in Section 2.5.1. Based on this correlation, hereafter we refer to a website as being conformant to a version of WCAG at a particular level if its home page reached a level of conformance.

The audits of the homepages of the websites from the DRC study established that only 11 out of 72 websites (about 15%) achieved level A, the minimum level of conformance with WCAG 1.0. This sample would not be comprehensive enough to represent the possible different levels of conformance, as it would not allow for comparisons with the accessibility levels of websites at higher conformance levels.

A search was undertaken for websites that reached higher conformance levels. 400 websites were found through Google searches on website conformance claims. Of these, only 45 of these websites did not fail automatic testing on their home page. Full audits of the home revealed that only 5 of those websites were actually conformant to any level of WCAG. Unfortunately, of those 5 there were none that reached AAA conformance for WCAG 1.0 or WCAG 2.0.

Further this search, websites from an implementation report of WCAG 2.0 (Web Accessibility Initiative 2008) were included in the candidate websites sample. This report presents websites that were considered to be conformant with WCAG 2.0 at different levels, as supporting evidence from the working group. These websites were audited for both WCAG 1.0 and WCAG 2.0.

The following are the 16 websites selected for the study presented with the tasks that were used in the study. Each task is listed with the number of web page steps on the optimal path required to complete the task are listed in parentheses.

- www.lflegal.com: Law Office of Lainey Feingold is an office specialised in disability rights with long texts and legal jargon on the page
    - Find the definition of “structured negotiation” (3)
    - Find what the pharmacy chain Rite Aid has agreed to do regarding the use of captchas in their website (3)
    - Find the deadline for Staples to install tactile keypads on their point of sale system in their US stores (3)
- www.green-beast.com: Green Beast Design is a site for web designers with an embedded blog.
    - Find the price charged per hour to develop a website (2)
    - Find a quote of the Military Audiology Association about Green Beast (2)
    - Find the name of the author of the introductory video about accessibility mentioned in a blog article (3)
- www.york.gov.uk: The City of York council website. This site contains complex forms and data tables.
    - Find the cost of council tax for properties in band E for 2010/2011 (2)
    - Find if there is a Park and Ride bus service to the Designer Outlet and what is the return fare (4)
    - Find what is the nearest primary school in the area of a given postcode (6)

- www.nhsnss.org: The National Health Service for National Services for Scotland provides people with online health information. This site contains PDF reports.
    - Find the address of the Blood Donor Centre in Inverness (3)
    - Find out when the NHSNSS was launched in the institutional video (3)
    - Find what the Executive Office has done to fulfil its disability duty, in latest the Disability Equality Scheme report (7)
- www.copac.ac.uk: The Copac, National Academic and specialist library catalogue website with complex search forms.
    - Find the name and address of a library in York in the network of libraries (3)
    - Find the name of a library that has the Harry Potter book number 4 available (4)
    - Find name of research and development coordinator – staff member (3)
- www.theaa.com: The Automobile Association website providing customers with information about car travel and insurance which contains complex data tables.
    - Find the telephone contact number for car insurance enquiries (2)
    - Find different car insurance plans that cover damage caused by fire or theft (3)
    - Find educational information targeted for schools and colleges on how to pass driving tests quickly (6)
- www.dh.gov.uk: The UK Department of Health website with multimedia content
    - Find the name of the Member of Parliament in charge of Public Health (3)
    - Find information contained in video campaign for swine flu (5)
    - Find report of the Plain English Workshop that happened in March 2006 (6)

- www.digizen.org.uk: Digital Citizen is an educational website with dynamically generated content and an interactive Flash application.
    - Find a list of risks associated with the use of social networks (4)
    - Find statistics about the number of young people that claim to have been target of cyber bullying (3)
    - Create a digital avatar using the "Digicentral" service (8)
- www.jisc.ac.uk: The Joint Information Systems Committee (JISC) is an agency for technology development in the UK. Their website contains audio podcasts.
    - Find the deadline and budget limit for next invitation to tender (2)
    - Find the venue of the 2008 JISC-supported event "Rethinking the digital divide" (5)
    - Find specific information in podcast "The financing of higher education" (5)
- www.royalmail.com: The website for the Royal Mail UK post services. This website contains complex tables and forms.
    - Find the form to request a redelivery (4)
    - Find the weight limit for a large first class letter (4)
    - Find the price to send a parcel weighing 5Kg to Spain in 3 days (5)
- www.pret.co.uk: The website of the major restaurant chain Pret a Manger which contains a PDF menu.
    - Find the nearest Pret a Manger shop to the university post code (3)
    - Find nutritional information about the Classic super club sandwich (4)
    - Find the price of the Luxury sea food selection for delivery (4)
- www.tuc.org.uk: The website of the UK Trades Union Congress. This website has a large amount of information on it including multimedia videos.
    - Find the telephone contact number and the name of the general secretary of the National Unions of Teachers (2)
    - Find how long a parent can spend on adoption leave (4)
    - Find who was the speaker on the Economy and unemployment debate in the Congress 2009, who made a point about what were frontline services in the NHS (3)

- www.britishmuseum.org: The website of the British museum which has a substantial amount of multimedia content and images.
    - Find the price of a painting with the River Thames (3)
    - Find the room in which the Snettisham Hoard is displayed and if it is one of the museum's most treasured exhibits. (5)Find information contained in video about Hadrian's wall (Path until video provided to user) (6)
- www.nhsdirect.nhs.uk: The general health advice website for the National Health Service in the UK. This website has interactive features with client-side dynamically generated content.
    - Find the name of the condition related to blood clots in long haul flights and ways to prevent it (5)
    - Find the nearest walk-in centre to the university post code (5)
- www.ford.co.uk: The automobile and truck manufacturer Ford. This website allows users to check car prices through an interactive client-side application.
    - Find the nearest dealer to the university post code (3)
    - Find the cheapest used Ford Fiesta within 200 miles of the university post code (4)
    - Filter cars available according to budget, seats, doors and fuel (6)
- www.ticketmaster.co.uk: The major ticket seller for events worldwide. This website has an interactive ticket booking system which includes a variety of dynamic content as well as CAPTCHAs.
    - Change the default location of the website to the university post code (2)
    - Find the next Jazz/Blues music event in the next 14 days (3)
    - Buy a ticket at the Grand Circle for the next event at the Grand Opera House in York (6)

This set of websites use a variety of commonly used different technologies in their implementation. Besides common HTML (Hypertext Mark-up Language) and CSS (Cascading Style Sheets), there were websites that used blogs, complex forms, complex data tables, images, information in PDF (Portable Document Format), multimedia content (both audio and video), interactive functionalities with Flash, dynamically generated content with Javascript and "captchas"[18]. They also include

[18] Characters displayed distorted on purpose, used as a security measure to check that the page is being used by a human. Supposedly, the distorted characters cannot be recognised by computer algorithms

websites from a range of sectors, both public and private, including public and commercial services, local, regional and central government websites, and non-governmental organisations.

The results of the WCAG audits of the home pages of these websites is presented in Table 3.2. Of the 16 websites, 4 were conformant to WCAG 1.0 Level AA, 8 were conformant to WCAG 1.0 Level A and 4 websites were not conformant to any level of WCAG 1.0. No websites reached WCAG 1.0 AAA.

According to the audits of the websites with WCAG 2.0, the websites were distributed as following, according to their level of conformance with WCAG 2.0:

- 1 website conformant with WCAG 2.0 at level AAA
- 1 website conformant with WCAG 2.0 at level AA
- 2 websites conformant with WCAG 2.0 at level A
- 12 websites that were not conformant to WCAG 2.0

The number of websites in the selection conformant to WCAG 2.0 was lower than then number of WCAG 1.0 conformant websites mainly due to changes in the guidelines and changes in the priority levels. New SCs were added in WCAG 2.0, and some SCs that were previously at priority 2, were placed at level A – notably WCAG 1.0 checkpoint 3.2 – “Create documents that validate to published formal grammars” at level 2, relating to WCAG 2.0 SC 4.1.1 at level A. New SCs included requirements for error identification (SC 3.3.1) at level A and error suggestion (SC 3.3.3) at level AA. Some of the websites that were conformant to WCAG 1.0 at levels A and AA did not meet SCs 3.3.1 and 3.3.3, and some of the websites that were conformant to WCAG 1.0 at level A had failed at HTML validation tests, which deemed them to fail WCAG 2.0 level A. When examining the ways in which websites failed to conform, it was noted that 3 of the websites (York, NHSNSS and The AA) failed one SC a single time, specifically SC 3.3.1 (“error identification”), and otherwise conformed to Level A of WCAG 2.0. Therefore, in some analyses, which will be noted, these websites are classified as Level A conformant websites.

Users performed tasks with 16 websites. For each of those websites a selection of 2 – 3 tasks were created, aiming at covering different aspects of the websites. The tasks devised for the websites were representative of the typical tasks that users would carry out on each website. It was attempted to have shorter and less complex tasks being performed first, with increasing level of difficulty towards the end.

Table 3.2: WCAG 1.0 and WCAG 2.0 audit results for the home pages of the 16 websites used in the study.

Websites	Number of different checkpoints violated per WCAG 1.0 Priority Level (P)				Number of instances of violations of checkpoints per WCAG 1.0 Priority Level (P)				Conformance Level	Number of different WCAG 2.0 SCs violated by SC Level				Number of instances of violations of WCAG 2.0 SCs by SC Level				Conformance Level
	P1	P2	P3	Total	P1	P2	P3	Total		A	AA	AAA	Total	A	AA	AAA	Total	
www.lflegal.com	0	0	3	3	0	0	5	5	AA	0	0	0	0	0	0	0	0	AAA
www.green-beast.com	0	0	4	4	0	0	16	16	AA	0	0	3	3	0	0	9	9	AA
www.york.gov.uk	0	0	3	3	0	0	23	23	AA	1	2	2	5	1	3	3	7	Fail
www.nhsnss.org	0	0	6	6	0	0	30	30	AA	1	2	6	9	1	4	26	31	Fail
www.copac.ac.uk	0	2	6	8	0	2	19	21	A	0	1	2	3	0	2	4	6	A
www.theaa.com	0	5	4	9	0	8	72	80	A	1	4	4	9	1	28	29	58	Fail
www.dh.gov.uk	0	6	6	12	0	19	39	58	A	0	2	4	6	0	10	21	31	A
www.digizen.org.uk	0	10	5	15	0	27	23	50	A	3	2	7	12	13	12	21	46	Fail
www.jisc.ac.uk	0	5	4	9	0	32	36	68	A	3	3	7	13	185	10	21	216	Fail
www.royalmail.com	0	5	4	9	0	37	54	91	A	2	2	3	7	86	3	14	103	Fail
www.pret.co.uk	0	15	8	23	0	110	36	146	A	7	4	10	21	80	25	36	141	Fail
www.tuc.org.uk	0	11	5	16	0	176	8	184	A	6	3	8	17	51	18	28	97	Fail
www.britishmuseum.org	1	5	4	10	1	11	18	30	Fail	3	2	3	8	11	21	54	86	Fail
www.nhsdirect.nhs.uk	1	4	3	8	2	13	115	140	Fail	7	6	7	20	89	45	29	163	Fail
www.ford.co.uk	6	14	6	27	40	57	27	124	Fail	14	8	11	33	140	63	41	244	Fail
www.ticketmaster.co.uk	4	16	9	29	77	452	199	728	Fail	16	8	11	35	854	149	115	1118	Fail

3.3.2 Demographic questionnaires

A demographic questionnaire was used to collect basic information about participants. Questions included information about gender, age, native language, internet usage, computer experience, education level, employment status, use of assistive technologies and enhancements (improvements to a given resource, such as audio description, subtitles) for the Web, and information about their disability.

Participants reported their computer experience in in a scale from 1 (not at all) to 7 (extensive). They also rated their expertise with their assistive technology (if they used it) at the levels: beginner, intermediate, advanced or expert users.

A copy of the questionnaire with user information is available in Appendix C.

3.3.3 Equipment and software

The evaluations were performed using a personal computer running the Windows XP Operating System (Service Pack 3), with processor Intel Core 2 Duo 3 GHz, 3 GB of RAM memory, equipped with speakers, keyboard, a 15" LCD screen, a Logitech webcam and a 2-button mouse with scrollwheel.

Blind users had the choice of either the JAWS 10.0 screenreader or WindowEyes 7.11 screenreader. Partially sighted participants could use screen magnifiers ZoomText 9.1, Supernova 11 or Virtual Magnifier 3.22. Participants also could choose one of three web browsers: Internet Explorer 8.0, Firefox 3.5 and Google Chrome 15.

Recordings of the users' concurrent verbal protocol, their facial expressions and the desktop of the computer during the browsing tasks were taken using Morae 3.1. Morae was set to record keystrokes and mouse events.

Morae's screen-capturing mechanism was not compatible with ZoomText, Supernova and Virtual Magnifier. The screen capturing acted on a different layer than that on which the magnifying software programmes worked. This meant that Morae did not record the screen with the magnification and colour changes performed by those programmes. Recording the screen exactly as it was shown to the participant was very important for the analysis of the accessibility problems. Hence, a different set up was necessary for screen magnification users.

The set up for screen magnification users involved having the video output being sent to two monitors, by means of a signal splitter. One monitor was displayed to the user, and the other monitor was used to record the output using a camcorder. The recordings were performed using a Panasonic SDR-S15 digital camcorder.

3.4 Participants

3.4.1 Recruitment of participants

In order to reach participants from different groups, the following strategies were approached to recruit participants with print-disabilities:

- Invite participants from previous studies
- Advertisement at the Disability Support Service at York St. John University
- Advertisement at the Disability Support Service at the University of York
- Advertisement at the British Computer Association of the Blind (BCAB) mailing list
- Advertisement sent through the Action for Blind People mailing list
- Advertisement at the York University Students Union mailing list
- Invitation sent to the York Blind and Partially Sighted Society
- Advertisement at the Facebook group of the British Dyslexia Association

3.4.2 Description of blind participants

For this study, we considered blind participants as those who were totally blind, or whose residual vision was not enough for them to be able to see information in a computer monitor. The panel of blind users comprised of 32 participants, of whom 22 were male and 10 were female. Their ages ranged from 18 to 65 years (median = 39). Regarding their visual impairment, 17 participants had no residual vision, 12 had only light/dark perception and 3 had a very little central vision. Most participants (20 out of 32) had been blind since birth, and the remainder of the participants had had their condition for between 3 years and 47 years of age.

All participants used screenreaders as their primary assistive technology to access computers, 30 out of 32 use JAWS® and 2 use WindowEyes®. The WindowEyes® users used version 7.11. JAWS® versions varied from as early as JAWS 5.0 to JAWS®

11.0 (the most up-to-date version available when the study was conducted); 1 participant reported to use Jaws 5.0, 2 participants used Jaws 8.0, 3 participants used Jaws 9.0, 6 participants used Jaws 10.0 and 18 participants used Jaws 11.0. Regarding the participant using the oldest version of Jaws 5.0, he reported that he was comfortable to use Jaws 10.0, as he had already used this version previously, including in accessibility evaluations. Other participants who had older versions of Jaws at home also did not report difficulties using Jaws 10.0 during the tests.

Regarding their expertise with their screen reader, 10 users rated themselves experts, 9 as advanced users, 11 as intermediate and 2 as beginners. When asked to mention enhancements that they use, 17 participants reported to use audio description in multimedia content, and 6 participants reported to use text-only versions of websites if they are available.

In a scale from 1 (not at all) to 7 (extensive), participants' ratings of computer experience ranged from 4 to 7, with 87% of the participants rating their experience as 5 or above. Most of the participants (29 out of 32) had been using the Internet for 7 years or more. Internet Explorer was the most popular internet browser used by participants, being mentioned as primary navigator by all but one participant, who used Firefox. The majority of blind participants had English as their first language (30 out of 32). One participant had German as first language, and another had Gujarati, but both were fluent in English.

Regarding their education, 14 participants had completed secondary-level education, 2 had a trade qualification, 12 had a university degree, and 4 had completed post-graduate studies. With regards to their ability with Braille, 29 of the 32 participants reported to have some knowledge in Braille, with 20 rating them as experts in Braille, 4 as having advanced knowledge, 4 at intermediate and one at basic level.

3.4.3 Description of partially sighted participants

Partially sighted those participants were those who had some problems with their sight, but were not totally blind. The panel of partially sighted users comprised of 19 participants, of whom 9 were male and 10 were female. Their ages ranged from 21 to 68 years (median = 43). Nearly half of the participants (9 out of 19) had their sight condition since birth, and the remainder of the participants had had their condition for between 12 years, and 54 years of age. Only 8 out of 19 participants had previously done some evaluation of websites before.

Participants were asked to provide a description of their sight condition and how it affected the way they see and interact with computers. The conditions reported by participants were varied, and participants had different degrees of sight. The following is the description of each participant's sight.

- Participant 1 has different vision aptitude for the top and bottom of her eyes. She usually reads the web at a minimum font size of 16 point.
- Participant 2 is colour vision deficient. He has 20% of his/her central vision and 30-40% of his/her peripheral vision, and can only see black and white.
- Participant 3 can only see outlines and contours at or beyond a distance one meter.
- Participant 4 is 3/4 blind in her left eye.
- Participant 5 findings it difficult to see things on computers when colours are too bright or too dark. Participant 6 hasvery little sightwith only a little peripheral vision on his left eye, which enables to see black and white.
- Participant 7 has 4/60 vision acuity.
- Participant 8 has visual acuity 6/60.
- Participant 9 only has peripheral vision. She has to enlarge text to read, but too much enlargement is not adequate. She sees as if there was a line in the middle of the eye.
- Participant 10 lost his central vision on both eyes and only uses his peripheral vision.
- Participant 11 reported that her eyes have do not take light in. She has night blindness and tunnelled vision.
- Participant 12 has problems with involuntary movement of eyes and problems with depth perception.
- Participant 13 can see details at 3 metres with left eye, at 1 metre with right eye. She normally uses font size at a minimum of 24 point.
- Participant 14 has only 10% peripheral vision on both eyes and normal central vision.
- Participant 15 is totally blind in his right eye, and has distorted vision in his left eye. He has visual acuity 6/60.

- Participant 16 has little peripheral vision, reduced central vision and visual acuity 6/60.
- Participant 17 has tunnelled vision.
- Participant 18 can only see at very close distance.
- Participant 19 only has peripheral vision.

About three quarters of the participants (14 out of 19) used a specialised screen magnification software. ZoomText® was used by 9 participants, Supernova by 3 participants, and Lunar and Virtual Magnifier by one participant each. Out of these 14 participants, 4 reported to use speech synthesis as well as magnification when using computers. ZoomText® users used version 9.11, and Supernova versions ranged from 10.1 to 11.5. The Lunar® and Virtual Magnifier users did not know the version of their assistive technologies.

Two participants rated themselves as beginners, 7 as intermediate users, 2 as advanced users, and 3 as experts. The 5 other participants who did not use a specialised screen magnification software had different adaptations, including the use of screen resolution of 800x600px and extra-large scheme on Windows (used by 3 of the 5 participants), and resizing text on the internet browser (2 of the 5 participants).

The level of magnification used by each participant also varied considerably. Among the screen-magnification users, the level of magnification varied from 2 times to 36 times magnification, with 8 of the 14 participants using magnification levels between 2 times and 4 times. The participants who did not use screen magnification software used text zooming features in their browsers, with zoom levels between 120% and 200%. It is worth noting, though, that 3 of the 5 participants used zoom in their browser on top of a significantly enlarged screen with their Windows settings.

Regarding the colour settings, 9 of the 19 participants needed to change the colour scheme to be able to see the screen. Of these, 6 used an inverted colour scheme, one used yellow on blue, one used white on black and one used a monochromatic scheme with black on white. It is worth noting that 2 of the participants who use inverted colours did not use the feature provided by their screen magnification software. They preferred using the high contrast colour scheme from the operating system, as they said that when the magnification software inverted large sections of black content, the bright white resulting from the inversion of black would cause them a lot of discomfort in their eyes due to glare.

With respect to their experience with computers, in a scale from 1 (not at all) to 7 (extensive), the rating of experience ranged from 2 to 7; 74% of the participants (14 out of 19) rated their experience as 5 or above. Most of the participants (15 out of 19) had been using the Internet for 7 years or more. Internet Explorer was the most popular internet browser used by participants, being mentioned as primary navigator by 15 participants. Chrome was used by 2 participants, and Firefox and AOL were used by one participant each.

The time spent using websites varied from 1-5 hours per week to more than 20 hours. The majority of the users (16 out of 19) reported to spend more than 20 hours per week using websites.

It was not possible to have each website tested by ten different partially sighted participants. The target of seven different users per website was achieved, except for one website (NHS Direct) that went through a considerable overhaul and did not offer the same features as it did when the tasks were elaborated. However, having at least seven different users evaluating each website was still a good number of users and allowed for a detailed analysis of the problems they encountered.

3.4.4 Description of dyslexic participants

The panel of dyslexic users in this study had 13 participants, of whom 6 were male and 7 were female. Their ages ranged from 19 to 49 years (median = 20). The majority of the participants (12 out of 13) had English as their first language; one participant had Persian as first language, but was fluent in English. All participants had been diagnosed with dyslexia either by professionals linked to the University of York's Disability Office or by other external qualified professionals.

In a scale from 1 (not at all) to 7 (extensive), participants' ratings of computer experience ranged from 3 to 7, with 84% of the participants with experience rated as 5 or above. All the participants had been using the Internet for 7 years or more. The participants spent between 1 and 20 or more hours per week on websites; 6 out of 13 reported to spend more than 20 hours a week using websites.

Participants were asked to provide details about their dyslexia, in terms of how severe it was and in which difficulties they had associated with their dyslexia. Most participants reported to have been assigned a severity level in a severity scale that ranged from "mild", "moderate" and "severe", according to the results from psychological tests performed by specialists. In the sample of participants, 3 reported to have mild dyslexia, 3 mild-moderate dyslexia, 2 moderate dyslexia, 1 moderate-severe dyslexia, and 4 were not able to inform their level of dyslexia.

The difficulties reported associated with participants' dyslexia were very broad, and varied considerably from participant to participant. The issues reported and the numbers of participants affected by each of them are as follows:

- Difficulties with spelling (8 participants)
- Difficulties with reading and comprehension (7 participants)
- Difficulties with reading text with black printing on white background (7 participants)
- Limited short-term memory (4 participants)
- Low writing speed (2 participants)
- Difficulties with processing of verbal information (2 participants)

Some participants also reported issues that may co-occur with dyslexia such as difficulties with motor coordination (1 participant), limited spatial awareness (1 participant), speech difficulties (2 participants) and Asperger's syndrome (1 participant).

Five participants reported using some kind of assistive technology: 2 participants reported using Dragon Dictate and 2 participants use Dictaphone, both for speech recognition; 1 participant reported using TextHelp as a speech synthesizer software for reading texts on a computer. However, none of the participants requested to have these programmes installed for their tests. Regarding their enhancements, 6 participants reported that they normally change background colour of text in order to be able to read it comfortably, especially in word processors (however, none changed colour background on their web browsers during the evaluations), and 1 participant reported often increasing font size in websites to read text comfortably.

3.5 Procedure

The study sessions took place at the Human-Computer Interaction laboratory at the Department of Computer Science of the University of York. Participants were made comfortable as they arrived at the lab, and were briefed about the nature of the study, the process of the evaluation, the rating of problems when they were encountered, and the use of the concurrent verbal protocol. At this point, participants were asked if they had any questions, which were then answered by the researcher. After having any questions answered, participants were asked to read and sign an informed consent form. For blind participants, the researcher read out the informed consent for them. A copy of this document is presented at Appendix A.

Sessions lasted for up to two hours, and participants were offered breaks and refreshments during the sessions. Some participants who travelled from other places to take part in the study had more than one session during the day. In those cases, longer breaks were provided between sessions, as well as meals to participants.

Before the arrival of the participant, the laboratory was prepared according to the needs of the participants, including the special set up with two LCD screens and a camcorder for screen magnification users. A pre-defined configuration file was created with all recording settings for Morae. Any required assistive technology previously informed by the participants was also installed before the session.

After the participant signed the informed consent form, the participant was asked to start his/her assistive technology, if any was used. If an assistive technology was used, the participant was also asked to change any settings that he/she would like to change to use according to their preferences. For screen reader users, most changes included the speech rate, voice option, and other details in the screen readers. For screen magnification users, most changes were related to changing the level of magnification, the colour scheme, and switching speech on or off. Participants were also given some time to familiarise themselves with the computer, keyboard and other settings before starting the evaluation. At this moment, participants were also asked to state if they needed any adjustments in the physical environment, such as opening windows for ventilation, using a fan, or dimming the lights, particularly in the case of partially sighted participants.

Once all settings in the computer were done, the researcher would start the recording of the session, either on Morae or with the camcorder. The participants would then be asked to open the internet browser and open the first of the websites listed for the session.

For each website, the researcher would read the description of the task to the participant, and inform them that any information about the task could be asked at any time during the execution of the task. The participants would then carry out their tasks and stop to point out any problems that they encountered. The researcher also asked the participants to "think aloud" while performing their tasks, speaking their thoughts aloud and commenting about things they liked or disliked on the websites. When a problem was encountered, participants were asked to briefly describe what the nature of the problem was, and to provide a severity rating. Given the focus of the study on finding problems and their severity, users were also prompted to provide more information about problems they encountered when the researcher felt more detail was necessary. When participants became silent for an extended period of time during the sessions, the researcher would prompt them asking what they were thinking. Following are some of the main prompts used by the researcher during the sessions:

- *What are you thinking?* or *What are you trying to do?* – when participants became silent for an extended period of time
- *Could you explain more about this problem?* – when users pointed out a problem with statements such as "I don't like this", and the researcher felt that more detail was needed to understand the nature of the problem they encountered
- *How severe do you think this problem is?* – when users forgot to assign a severity to a problem they pointed out, the researcher reminded them to do so

In order to avoid the ratings being on the recording, participants were asked to not say their ratings verbally, but to either use their fingers under the table to indicate the number of their rating or to point to the rating on a sheet containing the severity ratings. The problem rating form was used to take note of each problem found by the participants.

For each problem, the researcher took note of the location where the problem was found, e.g. section of the website or part of the home page, a short description of the problem and the severity rating provided by the participant. A copy of the problem rating form is provided at Appendix B.

After each task was finished, the researcher asked participants to rate their perceived level of difficulty to complete the task, ranging from 1 – very easy to 5- very difficult. Also, the researcher asked the participant if he/she would like to report any other particular problem they found while performing the task that had not been previously reported.

Once all the tasks set for a website were finished, the researcher would ask participants to summarise the best and worst aspects of the website. During the sessions, participants would be given breaks and provided refreshments, depending on the length of the sessions. After the breaks, the procedure for evaluation of websites would be repeated for as many websites as there was enough time available to evaluate. After all websites were finished, the researcher would turn off the recording and save the video files.

After all websites had been evaluated, participants would be debriefed about the study, and given an opportunity to ask any questions they had about the study. Participants were then asked to sign the final section of the informed consent form, stating that they have been adequately debriefed, they have not been forced to complete the study and that all their questions have been answered.

At the end of the session, participants were asked to answer general demographic questions and specific questions about their disability and use of technology. After this, the researcher would collect any information that would be necessary for the compensation for the participation in the study, which could be e-mail or physical addresses for vouchers.

When participants came from different towns, all transport arrangements were made, and the researcher would also accompany participants to the rail station when necessary.

After the session, the researcher would run through the videos as soon as possible and archive the files with the main web pages visited by the participant in a given session.

3.6 Data analysis

The total amount video footage of the evaluation sessions was more than 200 hours of videos, being 100.9 hours of recordings of evaluations by blind participants, 45.8 hours by dyslexic participants and 65.7 hours by partially sighted participants.

This section describes the details about the activities performed to analyse these videos, including coding of user problems, mapping instances of user problems and distinct problems and matching user problems and relevant technical guidelines related to them.

Coding user problems involved two phases: the first phase, described in Section 3.6.1, consisted of the analysis of a subset of videos by three independent coders, in order to build a classification scheme with categories of problems that were mutually agreed. The second phase, described in Section 3.6.3, consisted of the coding of the entire set of videos.

The next activity following the coding of user problems was to identify *distinct problems* on websites and identify different *instances* when they occurred to different participants or to the a same participant in different occasions. This activity is described in Section 3.6.4.

Matching user problems and relevant WCAG 1.0 CPs and WCAG 2.0 SCs was the next activity, aimed at analysing whether the problems were covered by the technical guidelines. The process of matching problems and guidelines is described in Section 3.6.5.

Finally, the last activity performed in the data analysis was an audit of interface components in web pages that contained user problems with relevant CPs/SCs, as described in Section 3.6.6.

3.6.1 First phase of coding of user problems– definition of classification scheme

The aim of the first phase of coding was to build up a classification scheme for user problems, based on a mutually agreed set of categories established by the independent analysis by different coders.

A selection of videos for this first phase included users from the different disability groups and a range of different websites. Each video was initially coded independently by three different coders (the author and his two supervisors), who identified accessibility problems and assigned them an initial classification and severity rating.

After the independent coding of the videos, the three coders met to compare their initial identifications and classifications. During these meetings, a unified list of problems identified by all the coders was produced. Each problem was classified and a descriptive category was created. Based on the categories that emerged from these discussions, a classification scheme itself was built up.Table 3.3 shows an example of some problems compiled in one of the initial sessions to build the classification scheme

This initial phase of coding involved coding a set of 11 videos with blind, partially sighted and dyslexic participants on the British Museum, Lflegal and Ticketmaster websites.

Table 3.3. General list of problems identified by independent coders in the initial phase of coding

Problem No/Time	Identified by	Problem description from user/coder	Preliminary code	Type
P1 0:27	CP	"21 headings". User grimaced, resigned voice.	Heading Too many headings	Coder identified
P2 4:38	CP, HP, AF	"Heading level 2 highlights" .. I'm not really sure what highlights is.. what it's trying to indicate	Headings Heading content not meaningful	User rated
P3 5:27	CP, HP, AF	What it hasn't done is tell me where it is	Content: Expected content not on page	Coder identified

3.6.2 Categories of accessibility problems encountered by print-disabled users

The main goal of the categorisation scheme was to provide a description of the nature of the main problems encountered by print-disabled users when using websites. The categorisation was divided in problems that were related to six levels: *Content*, *Delivery media*, *Web page structure*, *Website navigation*, *Information Architecture* and *Underlying System characteristics*. Each category under each of the six levels contains *sub-categories* that describe the nature of the problem encountered by users. During the categorisation, each user problem was assigned to a sub-category. During the construction of the Table 3.4 shows a list of the categories defined under each of the levels and a description of each category.

Table 3.4. Description of top-level categories of user problems

Level	Category	Description
Content	Content (Meaning)	This category refers to the content in a web page, or the meaning the author wanted to convey
Delivery Media	Text	Issues related to the delivery of content as text
	Images	Issues related to the delivery of content as images (including pictures, graphs)
	Audio, Video and Multimedia	Issues related to the delivery of content as audio, video or multimedia (including animations)
	Other media types (music, mathematical notation, chemistry, etc)	Content delivered in media using other abstract notation
	All media types	Issues related to the delivery of content in any media type
Web page structure	Headings	Issues related to the use of headings and page structuring with headings
	Links	Issues related to individual link elements in a web page
	Tables	Content organised in table structures with rows and columns
	Controls, forms and functionality	Issues related to controls, form elements and functionality implemented in a web page
Website Navigation	Navigation	Issues related to the overall navigation as the structure for changing between pages in a website

Level	Category	Description
Information Architecture	Information architecture	Issues related to the way information is organised and structured in a website
Underlying System Characteristics	System characteristics	System issues related to the web application's underlying system, such as processing speed or to the assistive technology

The structure in the levels of *content* and *delivery media* was based on a conceptual framework for accessibility defined by Power et al. (2009). This conceptual framework makes a distinction between the *content* or *meaning* that is conveyed by means of a webpage and the *delivery media* that are used to encapsulate this content, be it text, images, audio, video or multimedia.

At the content level, subcategories included problems where content was not found where expected by users (missing content), problems where users could not make sense of content, irrelevant content before task content, illogical organisation of content, too much information in pages, difficult language of content and meaning in content lost due to transformations (such as text simplification).

The conceptual framework proposed by Power et al. (2009) also discusses *content adaptation rules* that must be applied to make content delivered in different media available to disabled people. According to this framework, these transformations produce either an *alternative* that replaces the original resource, or provide an *enhancement* to the original resource.

For example, content provided using an image as medium needs an *alternative* textual description of the content conveyed in the image in order to enable blind users to have access to this content; in this case, the alternative text replaces the image for these users. In the case of content conveyed using a video as medium, deaf users would need an *enhancement* in the form of subtitles or sign language interpretation to augment the original video.

Each category under the *delivery media* level had sub-categories describing problems related to alternatives, enhancements or presentation of content. The following sub-categories were included for the categories *text*, *images*, *audio, video and multimedia*, and *other media types*:

- **No alternative**: absence of alternative to a given resource, such as an alternative text to an image for a blind user.
- **Inadequate alternative**: alternative to a given resource is not adequate, such as an alternative text to an image for a blind person that does not describe the content of the image adequately.
- **No enhancement**: absence of an enhancement to a given resource, such as the absence of an audio-description to a video for a blind user or the absence of captions for a deaf user.
- **Inadequate enhancement**: enhancement to a given resource is not adequate, such as inaccurate captions in a video for a deaf user.
- **Default presentation not adequate**: the presentation of a resource is not adequate, such as text being too small for a partially sighted user, text that is not read out properly by a screen-reader or bad colour contrast in a visual medium.
- **Inability to change presentation**: presentation cannot be changed by specific settings in the user's browser or assistive technology, such as inability to change the size of a video on the screen.

Problems related to difficulties with scanning for content of any media type were in a category named "All media types", as it could be applicable to any media type.

The categorisation scheme also makes a distinction between problems that are related to an element contained *within* a single web page and problems that are related to the navigation *between* pages in a website. At the web page level were included the categories headings, links, tables and controls, forms and functionality.

The problems contained in the *links* category at web page level are to do with issues that are specific to a single link, such as having one link with unclear destination, link destination not present, poor link grouping, too many links or repeated links.

Problems related to the *Headings* category included issues such as having no or too many headings, headings that are not meaningful and illogical heading structure.

Problems related to the *Tables* category include problems where it was not possible to associate table cells to their headings, table structure being too complex, lack of headings and lack of alternatives to data in tables for users who find it difficult to handle tables.

Problems in the *Controls, forms and functionality* included issues with interface elements that could not be reached using a keyboard, lack of indication of how to interact with functionality, unclear description of what controls/form elements do, expected functionality not present, functionality not working as expected, no or insufficient feedback for actions, among others.

The website *navigation* level contains problems related to the navigation structure of a website and the issues that occur when *changing* between web pages. Such problems included navigation elements that do not help users find what they were seeking, no way to return to home page, navigation bar not salient, inconsistent navigation, destination not what anticipated by users, and impossibility to identify destination on arrival.

The *Information Architecture* level contains one category related to the organisation and structure of information in a website. This category includes problems such as complex organisation of content with too many steps to get to a web page.

The last level in the bottom of the categorisation scheme is *Underlying system characteristics*. Categories at this level involve issues that are related to system characteristics, such as issues with the web server where the application is being executed, broken links or system issues with the assistive technology.

3.6.3 Second phase of coding of user problems

The second phase of coding of the data was performed by one coder, the author of this thesis. In this phase, the remainder of the videos that were not coded in the initial phase were coded using the categories list built up during the initial phase. During this phase, if there was any new emerging problem that did not fit into the existing categories, a meeting involving three coders (the author and his two supervisors) would be set up to discuss the creation of new categories.

When coding the user problems encountered in the video analyses, both problems explicitly mentioned by the users and problems observed by the coder would be recorded. In cases where the problem was observed by the coder, the coder would also attribute a severity rating to the problem, using the four-point scale (cosmetic-catastrophic).

Each problem raised by a participant or observed by the coder was flagged in the respective Morae recording of the session. For each problem coded from the analysis of the sessions, the following information was recorded:

- **User problem**: a description of the problem as experienced by the user.
- **Technical problem**: a description of the technical causes of the problem.
- **Problem category**: the code assigned to the problem according to its category.
- **Severity rating**: severity rating assigned to the problem
- **Identified by user/coder**: identification of whether the problem was mentioned by the user or identified by the coder. This is especially important for the interpretation of the severity ratings of the problems.
- **Participant code:** code of the participant that experienced the problem.
- **Website:** website where the problem happened.
- **Task:** identification of the task that was being carried out when the problem occurred.
- **Web page**: identification of the archived web page where the problem was experienced.
- **Time when problem occurred** (generated by Morae)

The coding of user problems by the analysis of the videos was the most time-consuming activity in the research reported in this thesis. Excerpts of videos had to be watched several times in order to understand the nature of the problem that was being experienced by users and to identify possible causes.

The analysis was considerably harder for videos with blind participants using screen readers. It was crucial to understand what was conveyed by the speech synthesiser of those screen readers in order to know exactly what was happening when a blind user experienced a problem. Although it was asked that participants would use screen reader at a slower speed than that they would normally use, in many cases the speed was still too fast for the coder to follow. In these cases, it was necessary to set the presentation speed of the video to up to 0.7 of the normal speed. The time for coding one hour of video of blind participants could take at least three hours.

3.6.4 Matching instances of user problems onto distinct problems

More than one participant may have encountered problems caused by the same issue on a website, or users can have encountered instances of problems caused by the same issue more than one time. In those cases, it was important to identify different instances of problems that corresponded to the same issue on the website, that would be entitled a *distinct website problem*.

In order to determine that two or more instances of problems corresponded to the same distinct problem, two criteria were observed to establish that problems were the same:

- **The instances of user problems were of problems of the same nature**: In order to establish this, the problem category assigned in the coding of the user problems was used as a guideline.
- **User problems related to the same interface component**: Problems related to the same distinct problem had to be related to the same interface component on the same web page, or to an interface component that was repeated on several pages, such as problems with a fixed navigation menu.

Defining the set of distinct problems and problems that were related was very important to allow for a comparison between the severity ratings of similar user problems. It was also very important to observe particular problems on websites that would cause instances of problems more frequently.

The lack of clearly defined procedures for matching if different problem descriptions relate to the same problem or not can be a threat to the validity of studies that use such matchings (Hornbæk and Frøkjær 2008, Law and Hvannberg 2008). Investigating issues related to more reliable usability problems matching was one of the areas targeted at the project MAUSE (MAturation of Information Technology USability Evaluation) (Law et al. 2005). Law and Hvannberg (2008) described the process of matching problems (or consolidating problems) as involving the steps of problem extraction and problem filtering and merging, which can be done individually or collaboratively by coders. Hornbæk and Frøkjær (2008) conducted a study comparing different methods to establish whether different problems should be assigned as being the same. In their study, they found that matching problems by analysing the similarity of possible design changes in interface components to alleviate problems had the lowest agreement between evaluators when establishing whether problems were the same. The method for matching based on the User Action Framework (UAF) (Andre et al. 2001) had a better performance. Following this method, problems were categorised according to issues of related to the action cycle (Norman 1988) - planning, translation, physical actions, outcome and system functionality, assessment, and problems independent of the interaction cycle. Different instances of usability problems could only be matched if they were in the same category.

In the present study, the central point for the phases of filtering and merging problems to match different problem instances was the problem category assigned during the coding of problems, which was based mainly on the nature of the problems and how they affected users. By starting the matching of problems grouped by categories of user problems, the matching procedure relied on a categorisation of the nature of the problem, using the description of technical causes and related interface components only in a second step of the matching process. This method is more in line with the approach used by Hornbæk and Frøkjær (2008) based on the User Action Framework, which was more effective than approaches based on comparing possible design changes that would alleviate problems. In the UAF, problems were categorised according to the stages of Norman's action cycle, which is closer to the effect problems had on users' processing of tasks. The categorisation used in this work also aimed to portray how problems affected users.

3.6.5 Mapping user problems and technical guidelines

In order to establish whether WCAG 1.0 and WCAG 2.0 covered the problems experienced by print-disabled users, an analysis of each user problem was performed in order to match problems with CPs/SCs that could be relevant to each user problem.

The process of matching each user problem and CPs/SCs was performed in two stages. In the first stage, a search for possible relevant CPs/SCs was conducted, by means of a careful analysis of each CP or SC and its related documentation. In a second stage, when one or more CP/SC was found, a careful analysis of the nature of the CPs/SCs and documentation was performed in order to establish whether CPs/SCs were *directly relevant* to the user problem in question. A set of guidelines *covered* a user problem only if one or more CPs/SCs was identified to be *directly relevant* to the user problem, meaning it was clear that it addressed the problem encountered by the user.

There were cases in which the nature of a given CP/SC seemed clearly to be relevant to a user problem, but the CP/SC explicitly ruled out a certain case by including it in an exception. In these cases, the CP/SC would *not* be included as relevant to the user problem. For example, WCAG 2.0 SC 1.4.7 requires "Low or No Background Audio", but explicitly rules out this requirement for audio captchas. For this reason, in cases where users encountered problems with noise in audio captchas, SC 1.4.7 would not be considered as covering those problems. In order to avoid subjective interpretation as to whether user problems were covered or not by the guidelines, three accessibility specialists jointly analysed the initial matching of CPs/SCs with a selection of user problems from the most frequent problem categories that accounted for approximately 40% of all user problems. The rules for this classification were defined after consensus was reached among the three specialists during the discussions. Besides the most frequent categories of problems, other problems from other categories were also discussed between the specialists, especially when it was not clear whether a set of CPs/SCs were covered the nature of a given user problem in its entirety.

3.6.6 Audit of interface components connected to user problems with relevant WCAG 1.0 checkpoints or WCAG 2.0 success criteria

It was important to establish whether WCAG 1.0 CPs and WCAG 2.0 SCs related to a user problem were implemented or not by a given website. When one or more relevant WCAG 1.0 CPs or WCAG 2.0 SCs related to a user problem were found, the interface components connected to the user problems were audited using the methods described in Section 3.2.2 (for WCAG 1.0) and Section 3.2.3 (for WCAG 2.0). For each user problem with related CPs/SCs, specific interface components in the archived page visited by the users were audited.

The audits would be performed following the instructions in the WCAG 1.0 and WCAG 2.0 as closely as possible. In WCAG 2.0, for example, particular attention was paid to the descriptions of test procedures of techniques. Careful analysis of examples provided in the WCAG documentation also helped to determine whether the web pages analysed successfully implemented relevant CPs/SCs.

Audits of the related interface components could establish that CPs/SCs had been implemented or not. CPs/SCs were deemed as *implemented* if all the requirements for CPs/SCs relevant to a user problem were successfully implemented. There were cases in which two different SCs were related to a user problem, for example, when a user problem related to a link being unclear, two WCAG 2.0 SC were relevant to the problem – SC 2.4.4 – “Link Purpose (In Context)” at level A and SC 2.4.9 – “Link Purpose (Link

Only)" at level AAA. If a link in a web page passed SC 2.4.4 but did not pass SC 2.4.9, it would be counted as having successfully implemented one SC.

In some cases of audits of links, the assessment of whether the implementation of a CP/SC was successful may be subject to interpretation. When there was a user problem related to such cases, the outcome of the outcome of the audit would be assigned as "not implemented". For example, when a link text on a navigation menu is displayed on its own, such as a link "Explore" on a museum website. In this case, given that a user problem was reported with the link, the outcome of the audit would be assigned as the SC/CP not being implemented successfully by the website.

Chapter 4. User-based measures of the accessibility of websites and problems encountered by print-disabled users

This chapter presents the main results related to the primary research question proposed in this research: "*What are the main characteristics of accessibility problems encountered by print-disabled users when attempting to use websites?*" and related sub-questions. Section 4.1 presents results from measures of the accessibility of websites by evaluation with print-disabled users, describing task completion rates, task difficulty ratings, total instances of user problems and instances of user problems per problem category.

Sections 4.2, 4.3 and 4.4 present the characterisation of accessibility problems encountered by blind, partially sighted and dyslexic users, respectively. Each of those sections presents a characterisation of the problems encountered by each user group, with a list of the most critical problems in terms of frequency and severity.

A complete description of further types of problems encountered by each user group is presented in Appendix D. The description contains explanations of the nature of problems from the users' perspective and the main technical causes of those problems.

Section 4.5 presents an analysis of distinct website problems encountered by different user groups, including problems that were common to all user groups and problems that were specific to individual user groups. Section 4.6 presents a comparison between severity ratings of problems that were common to more than one user group and the severity of problems encountered by users of the same user group.

Section 4.7 presents a summary of the chapter.

4.1 Measures of the accessibility of websites evaluated by print-disabled users

Different measurements of the accessibility of websites were analysed in the present study. Part 171 of ISO 9241 (International Standards Organization 2008) defines as an extension of the concept the definition of usability Part 11 of ISO 9241 (International Standards Organization 1998) to “people with the widest range of capabilities”, including disabled people. Measures of usability as defined in ISO 9241-11 were used to understand the extent to which disabled users can use websites with efficiency, effectiveness and satisfaction. This section presents the results from the three user groups included in the study, regarding the *task success rates* (in Section 4.1.1), *task difficulty ratings* (in Section 4.1.2) and *instances of user problems* (in Section 4.1.3).

4.1.1 Task success rates

The first observed measure regarding users’ performance was whether they successfully completed the tasks attempted or not. Each of the 16 websites in this study was evaluated by between 10 and 11 different blind users, drawn from the sample of 32 blind participants. This yielded 478 tasks on websites attempted by blind users. Due to technical problems with the websites when the tasks were attempted, 17 tasks were removed from the analysis[19]. This resulted in 461 tasks considered in the analyses. Each website was evaluated by between 7 and 10 different partially sighted users[20], yielding to 322 tasks considered in the study. Each website was also evaluated by 10 different dyslexic users yielding to 468 tasks considered in the analyses.

---

[19] Task attempts not fully performed due to technical problems: Task 1 on www.britishmuseum.org by Participant 1, Task 2 on www.dh.gov.uk by Participant 8, Task 3 on www.royalmail.com by Participant 23 , Tasks 2 and 3 on www.tuc.org.uk by Participant 33, Task 3 on www.jisc.ac.uk by Participant 30 and Task 2 on www.digizen.org, Tasks 1, 2 and 3 on the www.ford.co.uk website and Tasks 1, 2 and 3 on the www.lflegal.com website by Participant 49 and Task 3 on the www.nhsnss.org and Tasks 1, 2 and 3 on the www.pret.co.uk website by Participant 67.

[20] The NHSDirect website was only evaluated by 4 partially sighted users, as it went through major changes that made it impossible to proceed with the evaluations.

Table 4.1 presents the task success rates for blind, partially sighted and dyslexic users, with the percentage of tasks succeeded and failed for each group.

Table 4.1. Task success rates for different user groups

User group	Tasks succeeded (%)	Tasks failed (%)
Blind	55.96	44.04
Partially sighted	48.99	51.01
Dyslexic	84.96	15.04

It is possible to observe from Table 4.1 that blind and partially sighted users had a higher percentage of tasks failed than dyslexic users. An Independent-Samples Kruskal-Wallis test on the success rates per participant showed a significant difference between the three user groups (KW = 23.64, df = 2, $p < 0.001$). A follow-up Mann Whitney test on the success rates of blind and partially sighted users showed no significant difference (MW = 229.00, n.s.). This showed that blind and partially sighted users had more problems to succeed in their tasks than dyslexic users, but had no significant difference in their task success rates.

4.1.2 Task difficulty ratings

A difficulty level was assigned by all users after attempting to perform a task on the websites. The difficulty was assigned in a scale from 1 to 5, where 1 was “very easy”, 3 was “neither easy nor difficult” and 5 was “very difficult”. Users were asked to rate the difficulty to perform tasks independently of having successfully completed each task or not.

Table 4.2 shows the mean task difficulty and the standard deviation of tasks performed by blind, partially sighted and dyslexic users.

Table 4.2. Task difficulty ratings for different user groups

User group	Mean Task Difficulty (1 – Very Easy / 5 – Very Difficult)	SD
Blind	2.84	0.605
Partially sighted	3.11	0.489
Dyslexic	2.38	0.26

An Independent-Samples Kruskal-Wallis test on the mean ratings of difficulty to perform tasks by participant showed a significant difference between the user groups (KW = 13.428, df=2, $p < 0.001$). A follow-up Mann Whitney test on the task difficulty ratings of blind and partially sighted users showed no significant difference (MW = 386.00, n.s.).

4.1.3 Instances of user problems

The number of problems encountered by users when attempting to use a website was an important measure of accessibility. A total of 3,012 instances of problems were encountered in the present study, being 1,383 by blind, 936 by partially sighted and 693 by dyslexic users. Table 4.3 presents the mean number of instances of user problems per website per user for blind, partially sighted and dyslexic users.

Table 4.3. Instances of user problems per user group

User group	Total Number of User Problems	Mean Number of Instances of User Problems per Website per Participant	SD
Blind	1383	9.22	5.31
Partially sighted	936	8.09	3.18
Dyslexic	693	4.64	1.33

An Independent-Samples Kruskal-Wallis test on the mean number of problems per website per participant showed a significant difference between the user groups (KW = 18.711, df=2, $p < 0.001$). A follow-up Mann Whitney test on the on the number of problems per website per blind and partially sighted users showed no significant difference (MW = 281.00, n.s.). Significant differences were found by Mann Whitney tests between the number of problems per website per blind and dyslexic participants (MW = 38.00, $p < 0.0001$), and between partially sighted and dyslexic participants (MW = 127, $p < 0.001$).

An analysis of the number of instances of user problems per user group was also performed concerning the distribution of problems in the categories of problems defined in this study, as described in Section 3.6.2. Table 4.4 presents the number of instances and the percentage accounted for problems in each of the top-level categories for blind, partially sighted and dyslexic users.

Table 4.4. Categories of user problems and frequency of instances for all user groups

Category	Blind N (%)	Partially Sighted N (%)	Dyslexic N (%)
Level: Content (Meaning)			
Content	324 (23.4)	214 (22.9)	241 (34.8)
Level: Media			
All media types	18 (1.3)	44 (4.7)	72 (10.4)
Text	26 (1.9)	186 (19.9)	54 (7.8)
Images	42 (3.1)	32 (3.4)	1 (0.1)
Audio, Video and Multimedia	62 (4.5)	56 (6)	17 (2.5)
Level: Webpage Structure			
Headings	111 (8)	1 (0.1)	1 (0.1)
Tables	25 (1.8)	10 (1.1)	1 (0.1)
Links	151 (10.9)	24 (2.5)	21 (3)
Controls, forms and functionality	364 (26.3)	195 (20.8)	141 (20.4)
Level: Website Navigation			
Navigation	208 (15)	147 (15.7)	127 (18.3)
Level: Information Architecture			
Information Architecture	15 (1.1)	12 (1.3)	8 (1.2)
Level: System characteristics			
Underlying System's characteristics	37 (2.7)	15 (1.6)	9 (1.3)

As can be observed in Table 4.4, the category with most problems encountered by blind users was "*controls, forms and functionality*", with 26.3%. Other categories also accounted for a substantial percentage of problems encountered by blind users, such as "*content",* with 23.4%, "*navigation"* with 15% and "*links"* with 10.9% of all the problems.

For partially sighted users, the category with most problems was "*content*", with 22.9% of all problems encountered by partially sighted users. Other categories that also accounted for a substantial percentage of problems encountered by partially sighted users were "*controls, forms and functionality*", with 20.8%, *"text"* with 19.9% and *"navigation"* with 15.7% of all problems encountered by partially sighted users.

For dyslexic users, the category with most problems was "*content*", with 34.8% of all problems encountered by dyslexic users. Other categories that also accounted for a substantial percentage of problems encountered by partially sighted users were "*controls, forms and functionality*", with 20.4% and *"navigation"* with 18.3% of all problems encountered by dyslexic users.

A Related-samples Friedman's one-way analysis of ranks showed no significant difference between the distribution of problems in the categories between the three user groups (X^2 = 0.522, N=12, df=2, p = 0.770).

4.2 Accessibility problems encountered by blind users

Blind users encountered 1,383 problems in this study. Out of those problems, 847 (61.2%) were reported and rated by users, 297 (21.5%) were mentioned by users but not rated and 239 (17.3%) were identified by the researcher. A total of 64 subcategories had instances of problems encountered by blind users.

This section presents a summary of the most frequent and most severe problems encountered by blind users.

4.2.1 Main problems encountered by blind users

From the 64 subcategories of problems encountered by blind users, the 15 most frequent subcategories accounted for 67.5% of the problems. Table 4.5 presents the list of subcategories, followed by the number of instances of problems that occurred and the percentage of the total number of problems accounted by each individual subcategory. The third column presents the median severity of these problems.

Table 4.5. Median severity rating of most frequent subcategories of problems encountered by blind users

Subcategory description	Instances N (%)	Mean Severity Rating
1. Link destination not clear (Links)	117 (8.46)	2
2. Navigation elements do not help users find what they are seeking (Navigation)	99 (7.16)	**3**
3. Content not found in pages where expected by users (Content)	88 (6.36)	2
4. Irrelevant content before task content (Content)	87 (6.29)	2
5. It is not clear what particular controls or form elements do (Controls, forms and functionality)	79 (5.71)	**3**
6. No/insufficient feedback to inform that action has had an effect (Controls, forms and functionality)	72 (5.21)	**3**
7. Users cannot make sense of content (Content)	66 (4.77)	2
8. Destination not what was anticipated (Navigation)	48 (3.47)	**3**
9. Functionality does not work (as expected) (Controls, forms and functionality)	48 (3.47)	2
10. Control or form element cannot be reached using the keyboard (Controls, forms and functionality)	44 (3.18)	**4**
11. No headings (Headings)	41 (2.96)	2
12. Organisation of content is inconsistent with web conventions/common sense (Content)	39 (2.82)	2
13. Inadequate alternative to image (Images)	33 (2.39)	2
14. No enhancement to audio, video or multimedia (Audio, video or multimedia)	31 (2.24)	**4**
15. Expected functionality not present (Controls, forms and functionality)	31 (2.24)	**3**

From the 15 most frequent subcategories of problems encountered by blind users, two had median severity rating 4 (catastrophe): "*control or form element cannot be reached using the keyboard*" and "*no enhancement to audio, video or multimedia*".

Other five subcategories had mean severity rating 3 (major): "*it is not clear what particular controls or form elements do*", "*navigation elements do not help users find what they are seeking*", "*no/insufficient feedback to inform that action has had an effect*", "*functionality does not work (as expected)*" and "*expected functionality not present*". The remainder eight of the 15 most frequent subcategories of problems encountered by blind users had median severity rating 2 (minor).

4.2.1.1 Control or form element cannot be reached using the keyboard

This subcategory was listed in the 15 most frequent categories of problems encountered by blind users with 44 instances, accounting for 3.18% of all problems encountered by this user group. Another important aspect of this category is that it had median severity rating 4 – catastrophe, meaning that at least 50% of its problems had the highest possible severity rating.

Problems in this subcategory occurred when blind users were unable to have access to a control or form element using the keyboard. In many cases, for example, users expected that there should be a button somewhere when they detected that a form had ended or when they were aware of the existence of an interactive component on the screen, but were not able to get access to the element.

In one example, users were looking for a video in a governmental website. Users went up and down in the page using the keyboard, going past the place that elements in the page seemed to suggest where the play button would be located, but they could not reach any button. In another example, users were trying to refine the search for a car in a vehicle manufacturer's website by budget. They found a text informing where they could select the "budget", but did not have access to the budget selector.

Regarding the possible technical causes of those problems, in most of the cases where this problem occurred, controls or form elements were not implemented accordingly to allow access via keyboard. Examples included Flash buttons that could only be activated using a mouse, such as in the cases of embedded videos that could not be played using the keyboard only or cases where controls were implemented using JavaScript that only allowed access using a mouse, such as in the example illustrated in Figure 4.1.

Figure 4.1 – Example of control – budget selector in a car manufacturer's website - that is not reachable using a keyboard

The criticality of this type of problem brings very important implications to design. It is very important that designers consider carefully the ways different users interact with websites. If interface components cannot be reached using the keyboard, blind users will not be able to use them at all, unless they try to use a mouse simulator in screen readers, which is used by very few users and with a difficulty that makes it impractical for users to use.

4.2.1.2 No enhancement to audio, video or multimedia

The lack of enhancements to audio, video or multimedia was also one of the most critical categories, also listed at the 15 most frequent with 31 instances and accounting for 2.24% of all problems encountered by blind users, and with median severity hitting the highest severity level.

Blind users needed to find specific information in audio, video or multimedia, but were unable to get all the information they needed due to the lack of an enhancement, such as audio-description. In one example, users had to find information in a video in a museum website combining what was contained in audio and information that was only shown visually on the screen. Due to the lack of audio-description of graphical information, they were not able to obtain all information involved in their task. Figure 4.2 illustrates the example of a scene of a video embedded on a website, describing the site of Hadrian's wall in Northumberland. The text in the captions with the speeches illustrates what blind users are able to know about the video. As there is no audio description of the visual scenery, those users lose very important descriptions of visual to fully understand this educational video, including the description of the area where the remains of Hadrian's wall are.

Figure 4.2 – Example of video without audio description – blind users are able to listen to speeches but are not aware of important visual information of scenery important to understand the scene

The importance of having audio description of video content for blind users is a very relevant finding of this work. In videos without audio description, blind users can miss a lot of information that is presented visually, such as scenes without dialogs, who are the characters speaking on certain scenes, and other important information that are essential to understand the message contained in videos.

Providing audio description can incur in extra cost for content producers, and it can also demand professionals with special expertise in producing audio description. However, the benefits brought to blind and partially sighted users is substantial, and this should be taken into account by content producers when making decisions about producing audio description of their material.

4.2.1.3 Navigation elements do not help users find what they are seeking

This subcategory had one of the highest frequencies in the types of problems encountered by blind users with 99 instances, accounting for 7.16% of all problems encountered by blind users. Besides, it had a high median severity rating 3 (major), meaning that at least 50% of the problems in this subcategory had severity rating major or catastrophic.

Users found that the navigation elements were confusing and disorienting, and did not help them find the information they were seeking in their task. In one example of problem in this sub-category, users were seeking the name of a cabinet minister in charge of public health in the Department of Health. The navigation had several options that seemed to be plausible, such as “Public Health”, “About us”, “Contact”, but users could not be sure which one to follow. Figure 4.3 shows the navigation bar of the Department of Health website with the options available. In this example, the information about the referred minister was under About us/ Ministers.

Figure 4.3 – Example of navigation of the Department of Health website – users had difficulties finding where to find the cabinet minister in charge of Public Health

The feeling of being lost was especially severe to blind users, as checking different possible options in navigation was very time-consuming due to the time taken by screen readers to read the content. This suggests that having clear navigation mechanisms can be especially important to blind users.

In most cases when problems in this subcategory occurred, there were not specific problems with individual links available in the navigation. The causes of these problems were strongly related to poorly designed information architecture in websites.

In order to solve those problems caused by poor information architecture, developers should devote special attention to designing the way information is organised in their websites. The first important aspect to design good information architecture is to strive to use descriptive labels in navigation structures, organised in a way that enables users to clearly identify which path to choose to arrive at the content that they need. A second important aspect is to make the distribution of content into different web pages in a way that pages have a coherent set of information in them.

4.2.1.4 It is not clear what particular controls or form elements do

Problems related to the lack of labelling of controls and form elements were frequent and severe, showing the importance of providing accessible descriptions of interface elements that provide functionality to users. Problems in this subcategory had 79 instances, accounting for 5.79% of all problems, and had median severity 3 (major).

Users encountered form elements or controls and could not determine what they would do. Examples include cases where users encountered buttons that read “unlabelled 1”, or form fields that had labels that were not meaningful to users, such as “A-Z”, or even form fields that gave users no label at all.

The main cause for problems in this subcategory was the use of unclear labels to identify controls and form elements, or the lack of labels or identification of those elements. The problems occurred with several types of elements, including HTML form elements, such as input fields, combo boxes, check boxes or buttons, and also with other interactive technologies such as Flash buttons.

If users cannot identify what an interface component does, this has a severe impact on blind users’ interaction, including not knowing what to input in a text field, or not knowing what a button does. Besides, many users sometimes have to use trial-and-error approaches to identify which of a set of possible unlabelled controls could perform the feature they want, with potential disastrous outcomes in their tasks.

In the case of HTML elements, the causes of many problems were related to the lack of a properly defined <*label*> element explaining the purpose of <*input*> elements, or <*label*> elements that did not explain the purpose of <*input*> elements properly. With components that used Flash technologies, many components had descriptors that were left with pre-defined values such as “unlabelled 1”, “unlabelled 2”, etc. Figure 4.4 shows an example from the JISC website. Users were trying to listen to a podcast, but their screen readers only identified a series of three buttons labelled “unlabelled button” 1, 2, and 3. This way, users could not tell which was the play button.

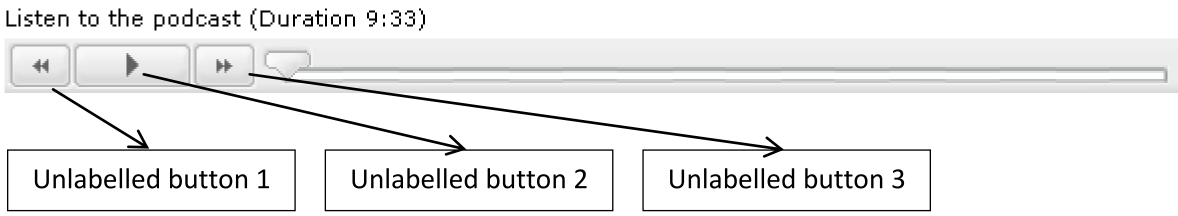


Figure 4.4 – Example of unlabelled buttons in a podcast player on the JISC website – reward, play and fast forward buttons are identified by “unlabelled button”

4.2.1.5 *No/insufficient feedback to inform that action has had an effect*

Problems with no or insufficient feedback to inform about the effect of an action had 72 instances, accounting for 5.21% of all problems encountered by blind users, and also had median severity 3 (major).

Users performed an action on the website and could not identify any feedback that the action had been performed. Problems included situations in which users activated a button or a link, and did not have any feedback if the action had had any effect. In many of these cases, their screen reader remained silent after performing an action.

Other examples included cases where some message was given, but it was not sufficient for users to recognise that the action had been completed. For example, in a city council's website, users searched for local services based on their address given by house number, street name and postcode. In the next screen, users encountered the message "select address", followed by a list of addresses, in case there could be more than one address under the same number (in a block of flats, for example). When reading this message, users did not recognise this as an indication that their action of informing the address had been completed. Although this indication was included in the website, the title of the page and a large part of the beginning of the screen (which is read first by screen readers) remained largely the same, making the users believe that nothing had happened. Figure 4.5 shows the illustration of this screen, with the message "select matched address" very further down in the page.

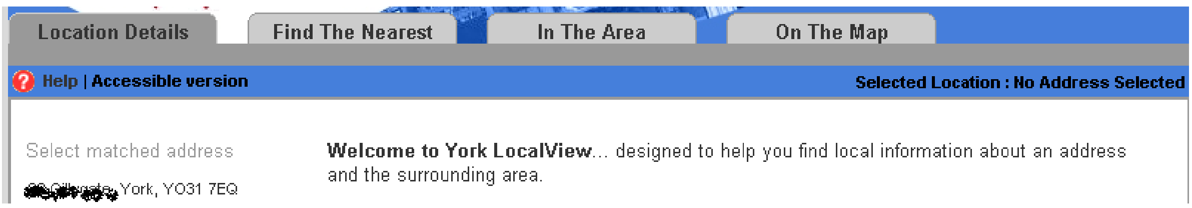


Figure 4.5 – Example of feedback about completion of action and direction for next step very further down in a page, where it took a long time for blind users to realise of the feedback

More than half of the problems in this subcategory were caused by the use of dynamic client-side features implemented on websites, such as features with Javascript or Flash. In one example, users activated a link named "Change location" in a ticket selling website, and the form to perform the action was included dynamically on the same page without reloading the current page on the browser. As this triggered no action on the browser, users did not know that anything had happened. Figure 4.6

illustrates the screen with the new content added dynamically on the page after activating the “change location” link on this website.

Figure 4.6 – Example of new content opened dynamically in the same page with no noticeable feedback to blind users

More than half of the problems encountered in the subcategory “*no/insufficient feedback to inform that action has had an effect*” encountered by blind users were related to content that was added dynamically to a page using Javascript without reloading a page. When such changes occurred, users were not informed about changes and, hence, did not know if they had had any feedback.

The use of technologies to include dynamic changes in pages has been increasing considerably, and it is of utmost importance that web developers and assistive technologies tackle accessibility issues related to it. The W3C is currently working on a set of recommendations for such applications, the Accessible Rich Internet Applications (ARIA) (Craig and Cooper 2011). However, this cited document is still under development, and improvements still need to be made to enable assistive technologies to work effectively with this kind of technology. Most of all, considerable research needs to be conducted with disabled users in order to design solutions to this issue that are effective to users.

4.2.2 Other important features of problems encountered by blind users

The problems encountered by blind users covered a wider range of types of problems than partially sighted and dyslexic users. Problems encountered by blind users were in 64 subcategories, while problems encountered by partially sighted users were in 54 categories, and in 43 for dyslexic users.

4.2.2.1 The importance of search features to blind users

Among the problems listed in the subcategory "*expected functionality not present*", it is noteworthy that a large amount of those problems were related to users expecting to have a search feature on websites. This is a very important consideration in favour of providing users with facilities to search for content on websites. The importance of search features to blind users can be justified by the considerable time it takes for them to browse through a website, due to the nature of their assistive technology based on speech. Being able to search for content instead of having to browser a website can significantly increase users' performance. However, it is important that search features work correctly and do cover all content available in websites. Many problems in the subcategory "*functionality does not work (as expected)*" corresponded to issues where search features failed to find content that users were confident was on the website using what seemed to be correct keywords.

4.2.2.2 Importance of headings to blind users

It was clear from this study that headings play an important role in navigation to blind users, with as many as 8% of problems encountered by blind users being related to headings. This is in agreement with the survey results from WebAIM (2011), in which 57.2% of 1,245 screen-reader users reported to use headings as a first approach to navigating within web pages. In the present study, not having any headings on pages was listed among the fifteen most frequent types of problems, but had median severity rating 2 (minor).

4.2.2.3 Identifying link destinations properly

Although not the most severe, problems with unclear destination of links were the most frequent type of accessibility problem encountered by blind users, with more than 8% of the problems. Blind users are affected by different problems from other users when they attempt to identify the destination of a link. In this study, many problems were related to link destinations being identified by poorly labelled images or by link destinations that do not make sense on their own, without considering the context in which they are placed.

Some accessibility guidelines have suggested that it would be acceptable to have link destinations that could be identifiable when placed in context (Caldwell et al. 2008), for example, a link identified as “read more” placed at the end of a paragraph with a summary of an article. However, more than 20% of the problems with unclear link destinations in the present study occurred with links that would make sense when placed in context. Many blind users used a feature in their screen readers that listed all links in a page, placed out of context.

4.2.2.4 Importance of content organisation to blind users

Inadequate organisation of content in web pages caused a substantial amount of problems to blind users in this study. As screen readers read content sequentially, the organisation of content can make it difficult for blind users to encounter the piece of information they wanted. The subcategory “irrelevant content before task content” alone accounted for more than 6% of all problems. Not all the problems were related to some classic web accessibility examples, where users find a navigation bar they have already read in previous pages before the content of a news article they want to read. Many problems were related to users finding advertisements and other irrelevant content that was not repeated in several pages before relevant content.

The subcategory “organisation of content is inconsistent with web conventions or common sense” also had a substantial number of problems. In problems in this subcategory, blind users also had to spend too much time scanning pages for content due to poor organisation of content.

4.2.2.5 Ineffective alternatives to CAPTCHAS and too little time to complete forms

It is well-known that CAPTCHAS are a serious problem to blind users, since the images with distorted text do not have associated textual alternatives on purpose, to avoid that computers recognise it. The common solution to visual CAPTCHAs is the use of “audio-captchas”, where a distorted audio contains a sequence of letters, numbers or words that have to be recognised.

One website in this study implemented an audio-captcha with sound in a noisy background. All users that attempted to use the audio-captcha found it extremely difficult to use it, and only three users managed to complete this task after a number of attempts that would be simply impractical in day-to-day activities. It is clear from the results in this study that simply providing “audio-captchas” with noisy background sound is not always an effective solution to providing an alternative to the visual captchas.

The few users who managed to solve this “audio-captcha” were faced with a message informing them that the system had timed out. They had been given only five minutes to try and solve the captcha.

4.2.2.6 Low severity of problems related to alternative texts to images

Assigning textual alternatives to images became a flagship and one of the most well-known issues of accessibility of websites to blind users. It was an unexpected result to find that those problems did not have a high severity rating from users. The subcategories related to this specific problems had median severity rating 2 (minor), including “*link destination is not clear*”, that had 36 problem instances related to poorly labelled images to identify a link destination. It could be that blind users may have become accustomed to encountering those problems and learned to ignore them.

Unless an image was crucial to users’ tasks, blind users tended not to rate those problems so severely. However, it is worth noting that this is a very frequent problem encountered by users, and it can become a serious annoyance to users, even when they are not necessarily interested in the images (such as in adverts), but have to listen to non-sense content from their screen readers.

It is also noteworthy that problems with images without proper labels were more frequent than problems with images with no textual alternatives at all.

4.2.2.7 Lower aversion to Flash

It seems that there are few problems with aversion to Flash from users. Although many specific problems were encountered in Flash applications, only 3 blind users reported problems where they had an aversion to Flash and wanted a separate alternative to it. That may indicate that some improvement has been made in the accessibility of Flash applications and its support from screen readers. However, it should also be noted that a substantial amount of specific problems were encountered with the lack of or poor labels of Flash elements, inability to use Flash interface components with the keyboard and other issues that should be addressed by developers when using Flash technologies.

Although few blind users had a total aversion to Flash, it is very important to highlight the importance of having special care to design interactive components that are usable by blind and other disabled users.

4.2.2.8 Dissatisfaction with text in PDF format

Unlike Flash, this study revealed that some blind users still have a dissatisfaction with text in PDF format, despite efforts to make files in this format more accessible. In 17 occasions, blind users reported problems with text being only available in PDF. Some of these problems occurred with PDF files that had accessibility features implemented. However, users still reported that they wanted an alternative in a different format. This could be due to repeated problems encountered previously by these users with PDFs, which made them have an aversion to files in this format. It could also be due to blind users still not being used to using different strategies to use PDF document from those they use with regular HTML content.

4.3 Accessibility problems encountered by partially sighted users

Partially sighted users encountered a total of 936 problems in this study. Out of those problems, 732 (78.2%) were reported and had their severity rated by users, 111 (11.9%) were mentioned by users but not rated and 93 (9.9%) were identified by the researcher. A total of 54 subcategories had instances of problems encountered by partially sighted users.

4.3.1 Main problems encountered by partially sighted users

Table 4.6 presents the 15 most frequently occurring problems encountered by partially sighted users.

The 15 of had mean severity 3 (major) or 4 (catastrophe), except for “*destination not what was anticipated*”, with median severity 2.5 (between minor and major). Three subcategories had median severity rating 4 (catastrophe): “*default presentation of* *image not adequate*", "*default presentation of control or form element not adequate*" and "*functionality does not work (as expected)*".

Table 4.6. Median severity rating of the fifteen most frequent subcategories of problems encountered by partially sighted users

Subcategory description	**Instances - N (%)**	**Mean Severity Rating**
1. Default presentation of text not adequate (Text)	157 (16.77)	3
2. Navigation elements do not help users find what they are seeking (Navigation)	78 (8.33)	3
3. Content not found in pages where expected by users (Content)	77 (8.23)	3
4. Difficult to scan page for specific item (All media types)	44 (4.7)	3
5. Default presentation of control or form element not adequate (Controls, forms or functionality)	43 (4.59)	**4**
6. Irrelevant content before task content (Content)	40 (4.27)	3
7. Expected functionality not present (Controls, forms or functionality)	33 (3.53)	3
8. Too much information on page (Content)	33 (3.53)	3
9. Default presentation of image not adequate (Images)	32 (3.42)	**4**
10. Inability to change presentation of audio, video or multimedia (Audio, video or multimedia)	31 (3.31)	3
11. Functionality does not work (as expected) (Controls, forms or functionality)	30 (3.21)	**4**
12. Organisation of content is inconsistent with web conventions/common sense (Content)	29 (3.1)	3
13. Users cannot make sense of content (Content)	28 (2.99)	3
14. Inability to change presentation of text (Text)	22 (2.35)	3
15. Destination not what was anticipated (Navigation)	22 (2.35)	2.5

4.3.1.1 Default presentation of control or form element not adequate

As one would expect, the presentation of graphical interface elements are very critical to the accessibility of websites to partially sighted users. Problems with the presentation of text, images, controls and form elements were listed among the most frequent and most severe. The presentation of controls or form elements, in particular, had the highest median severity – catastrophic in its 43 instances, which accounted for 4.59% of all problems encountered by partially sighted users.

Partially sighted users found it difficult or impossible to see or to interact with text or icons in controls or form elements due to inadequate presentation. In most cases, partially sighted users had to use different colour and size settings with assistive technologies or with settings in their operating systems.

The high severity of problems with the presentation of controls and form elements may be related to the criticality of those elements for users to complete their tasks. Not being able to see a button or what is in it can seriously hamper users' ability to complete tasks that would depend on them.

Problems in this subcategory were caused by poor colour contrast or small sizes of text and images in controls or form elements. Problems with colour contrast and size made it difficult for users to perceive where controls or form elements were or to identify their identification. Problems with small size and area of interaction also made it difficult for users to interact, since they required too much precision for users.

In one example, users had to select a seat to buy a ticket for an event by clicking on a circle on the seat map. The circle was very small, and some users complained that they "blended together". Most users also had problems to click on the circle due to its small size, which required a lot of precision from them. Figure 4.7 presents a screenshot of this seat map from the TicketMaster website.

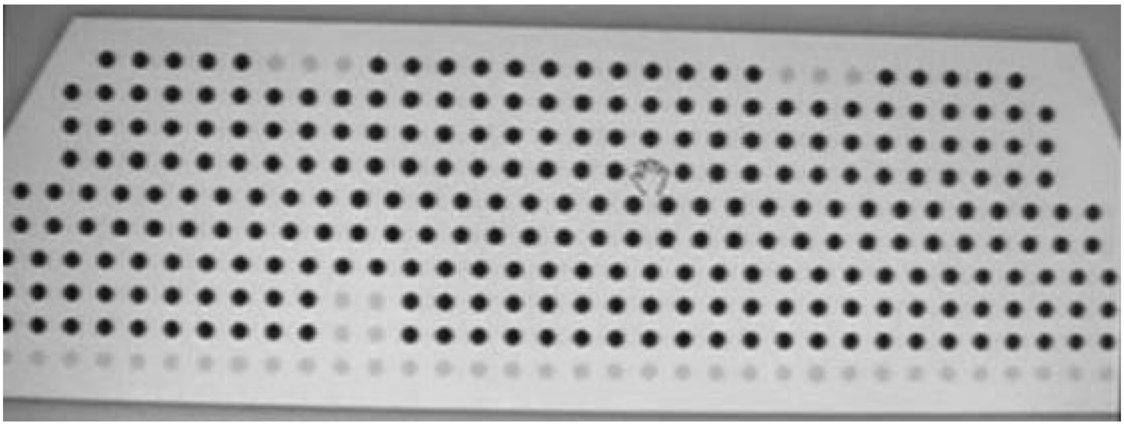


Figure 4.7 – Example of control with inadequate presentation – circles to select seat in a seat map have low colour contrast and are too small

Partially sighted users need to use a range of different adaptations for the visualisation of the screen, depending on their sight condition. Problems arise when those users encounter components with inadequate colour contrast, size or positioning. Interface components also need to be designed in a manner that allows for assistive technologies or settings in the operating system to perform changes effectively for partially sighted users.

4.3.1.2 Default presentation of image not adequate

Problems with the presentation of images also had one of the highest median severity ratings from partially sighted users – 4 catastrophe, and 32 instances, which accounted for 3.42% of all problems encountered by them.

Partially sighted users encountered problems seeing information in images when using screen magnifiers or when changing their settings to increase the size and change colour schemes. Users complained that images became blurred or pixelated when magnified. They also had problems when changes in colour background made it impossible to see what was in the image.

Another common issue encountered by users was the use of glary images on pages. Users found that images with bright white backgrounds gave them glare, often making them turn their faces due to the pain it caused or jeopardising their sight momentarily and preventing them from seeing what was shown around the glary image. In many cases, the image with glary white background would take a substantial part of the screen due to the magnification, which would increase even more the impact on users, who often had to use the computer when ambient lights switched off. Figure 4.8

presents an example of an image from the Pret A Manger website with a bright white background.

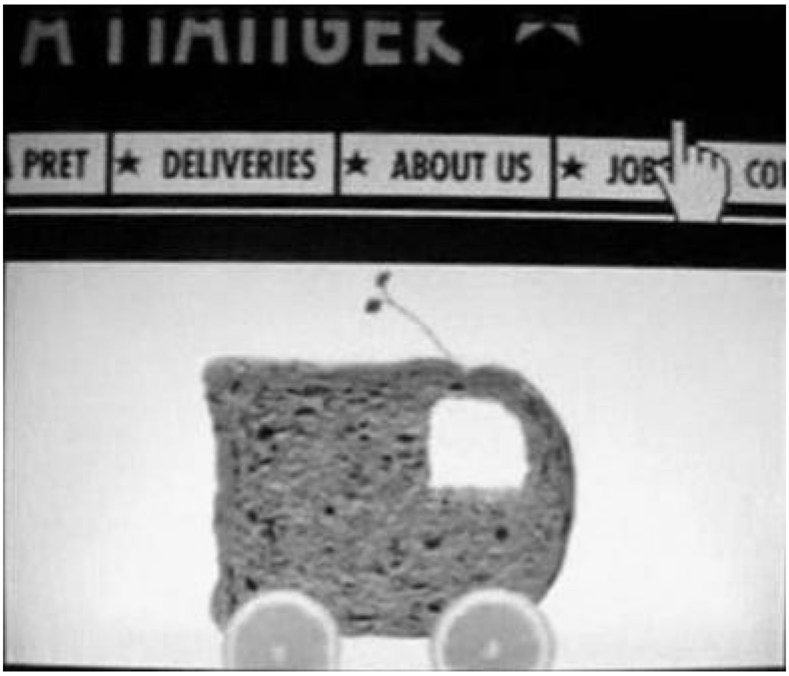


Figure 4.8 – Example of image with bright white background that causes glare to partially sighted users

Problems with glary images occurred mostly with users that had to change their colour scheme. When the change was made using the web browser's settings, changes in colour depended on the layout specifications in the web page. If an image has an opaque white background, its background will not change even if users choose to have a black background, as shown in Figure 4.8.

Images with low resolution are also the cause of problems with low readability of information in images. If the resolution is too low, images get pixelated when amplified.

In the case of images, a possible cause of the main problems encountered by users may be the lack of flexibility to change presentation settings in browsers and operating systems. For users who need special settings with all colours inverted, for example, specialised screen magnification software changes the colour of the entire final rendered presentation in the screen. When changes are performed via browser or operating system settings, on the other hand, the final result depends on how elements on the websites were designed. In the latter case, the results of changes depends on how interface elements were designed, and changes in images are considerably more limited than changes in text.

4.3.1.3 Default presentation of text not adequate

Although apparently easier to deal with technically than images, issues with presentation of text had a high median severity rating (3 – major) and the highest frequency of all subcategories encountered by partially sighted users (157), accounting for more than 16% of problems encountered by partially sighted users. This shows that it is urgent that considerably more attention be devoted to designing text with appropriate size, colour background and positioning.

Users encountered a substantial number of problems with text that they had difficulties to read due to inadequate presentation. In problems in this subcategory, users would somehow be able to change the presentation of text by changing colour contrast or the size. However, changing the settings would still result in a non-satisfactory presentation, or the necessary changes would demand too much effort from users.

The most frequent cause of problems with presentation of text, accounting for nearly 46% of problems in this subcategory, was poor colour contrast between text and its background. The colour used by web designers did not have enough contrast with the background in the default presentation, or became unreadable when users applied special colour settings.

Inadequate font size was another frequent cause of problems, accounting for approximately 18% of problems in this subcategory. Users complained that the font size was still too small even after magnification. In the example shown in Figure 4.9, the exhibition of a sandwich menu was very small for users, even though a very high level of magnification was used with the user's assistive technology. The problem, in this case, was due to the fact that the text was in PDF, and resizing had to be done on the PDF file reader as well as in the assistive technology, implying in more effort for users.

Figure 4.9 – Example of text of sandwich menu that is unreadable for partially sighted users, even after applying very high levels of magnification

Other cases include problems that occurred when resizing text implied in a disarrangement of the screen, with overlaying other text. In Figure 4.10, an example is shown where the price of an object in a shop is superposed by its description, which is not shown in its entirety on the screen.

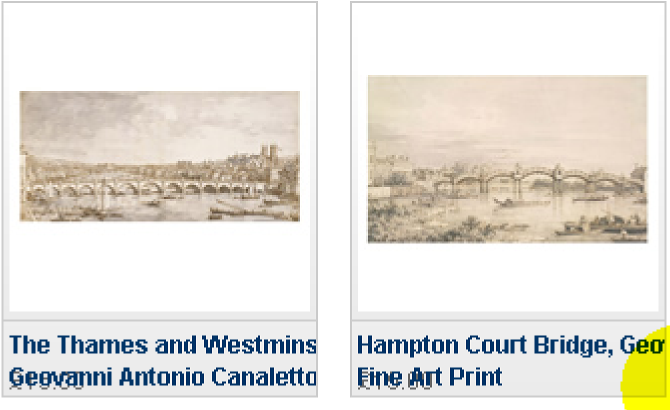


Figure 4.10 – Example of text that is not properly adjusted when enlarged and superposes other text

Other causes of problems also include use of text in images with low resolution that became blurred when magnified, or presentation of text in more than one column, that made users spend more time panning with their screen magnifiers.

4.3.1.4 Navigation elements do not help users find what they are seeking

Like for blind users, problems related to unhelpful navigation were critical to partially sighted users, with median severity rating 3 (major), and 78 instances, which accounted for 8.33% of all problems encountered by partially sighted users. Partially sighted users also found that the navigation elements were confusing and disorienting, and did not help them find the information they were seeking in their task.

Many partially sighted users spend a significant amount of time navigating in websites, especially when they needed to pan across different parts of the screen with high levels of magnification. Exploring different parts of a website by trial and error when navigation structures are not helpful can be very time consuming, and not finding content where expected can be very frustrating to those users.

In most cases when problems in this subcategory occurred, there were not specific problems with individual links available in the navigation. The reason for the problems was with the overall structure of the navigational elements, caused by a poorly designed information architecture. In one example of such problem, users were looking for a specific exhibit in a museum website. However, as shown in the illustration of the navigation bar in Figure 4.11, several different options seemed to be plausible to find this, such as “Visiting”, “What’s on”, “Explore” or “Research”. In fact, during the evaluations, many users had to look at several of these options by trial and error, which meant a very time-consuming process for users.

Figure 4.11 – Example of navigation in the British Museum website – users did not know what to follow to find a specific exhibit in the museum

4.3.1.5 Content not found in pages where expected

Partially sighted users encountered problems when they confidently followed a link to a page, but a piece of information that they expected to find there was missing. Those problems also had median severity 3 and 77 instances, accounting for 8.23% of the problems encountered by partially sighted users.

In most problems in this subcategory, users seemed to believe a given link was the right one to follow, but when they arrived at the page, it did not contain the content they wanted. The content was actually in another page different from that which many users believed was the right one. This was a very serious issue to partially sighted users. As many of those users had to pan around to see different parts of the screen due to their small viewport, they sometimes spent a lot of time double-checking the page they believe to have the content they believed to be there, just to make sure they had not missed the content.

In one example, users were looking for the names of staff members of the libraries network Copac. They found the link “Contact”, which they believed would have the list of contact names, but it only had the general contact information to the institution, as illustrated in Figure 4.12. The information about staff members was under “About”, not under Contact, as expected by users.

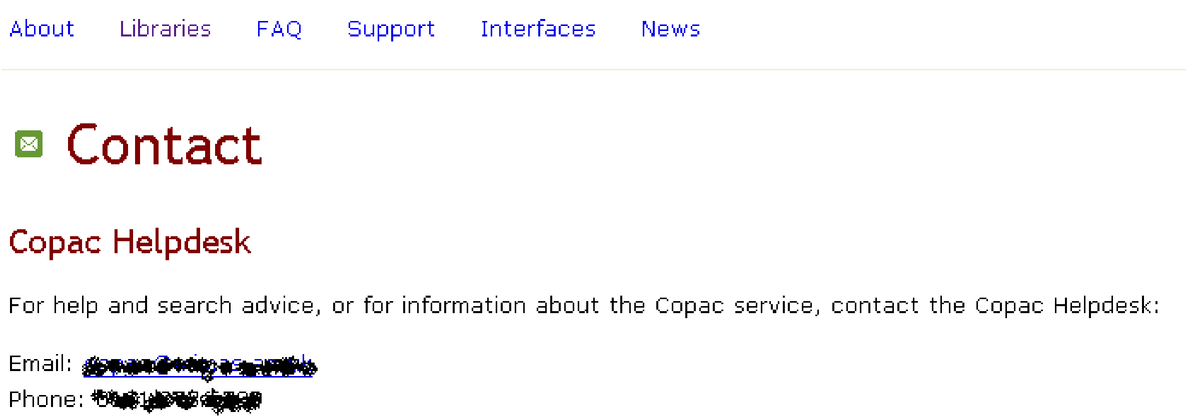


Figure 4.12 – Example of contact page from the Copac website that did not contain names of staff, as expected by users

Similarly to the problems where users found that the navigation did not help them find what they were seeking, the main technical cause of problems in this subcategory was connected to the information architecture and organisation of websites. Users expected to find certain pieces of information in parts of the website that were different from that laid-out by the designers of the websites.

4.3.2 Other important features of problems encountered by partially sighted users

4.3.2.1 High severity of problems encountered by partially sighted users

The severity of the problems encountered by partially sighted users is noteworthy. Different comparisons between the severity of problems encountered by partially sighted users and the other groups showed that the severity was higher for partially sighted users. Besides, out of the 15 most frequent subcategories of problems, 14 had median severity ratings 3 (major) or 4 (catastrophe), which was higher than the severity of the most frequent problems encountered by blind and dyslexic users.

4.3.2.2 Colour, glare and images in screen magnification software

Fourteen of 19 participants in the present study used specialised screen magnification software. Although this kind of software offers features to change the colour settings as well as size, some users reported to prefer using colour settings available on their web browser or operating system. The reason for this was that they like to see images with their original colour. When they use an inverted colour scheme, the entire rendered screen is inverted, including images.

A substantial number of problems were encountered with glary white parts of the screen that caused pain and jeopardised users' vision for a period of time. Those users normally needed to use darker background and use computers in a room with low levels of light. For users who did not use colour settings in specialised software, this normally happened because the colour background in large images and multimedia content did not follow the darker background settings they chose.

4.3.2.3 Scanning pages for content

Problems with difficulty to scan for specific items had high frequency and high severity ratings. This shows that it is very important that research be conducted to devise new approaches to help partially sighted users scan for content on the screen. Strategies to highlight important content and organise content in web pages may be developed. However, the particularities caused by the limited amount of information shown at a time with magnification software are particularly challenging in the search for new design and technological approaches to help those users in scanning tasks.

4.3.2.4 Organisation and amount of content in web pages

Specific issues related to the organisation and amount of content in web pages were very frequent, and also had high severity for partially sighted users. Three of those subcategories had median severity rating 3 (major) and together accounted for more than 10% of all problems encountered by partially sighted users: "*irrelevant content before task content*", "*organisation of content is inconsistent with web conventions/common sense" and "too much information on page*".

Those issues impact partially sighted users severely, as the operations to pan across parts of the screen are very time consuming. Users become frustrated when they have to go through a lot of irrelevant content placed before content of interest or when content is not organised logically. Pages with too much content also have an especially severe impact on those users, as it results in too much scrolling when different sizes are used.

4.3.2.5 Links that open in new windows

Many problems in the subcategory "destination was not what anticipated" were related to links that opened a new window without users' awareness. It has been widely acknowledged that blind users have problems when a new window is opened without their knowledge. However, the results in this study show that this issue also affects partially sighted users, especially those who use screen magnification software with high levels of magnification. Those users may find it difficult to recognise that a new window was opened, as they cannot see the full screen at one time. In the operating system Windows, for example, this recognition is normally done by acknowledging a new window in the bottom task bar. However, this part of this screen is not always necessarily shown in users' viewport.

4.3.2.6 Lack of responsiveness of AT with graphical elements

The subcategory "*assistive technology becomes irresponsive with particular graphic elements*" was exclusive to partially sighted users. However, there were occasions where blind screen readers used by blind users also became irresponsive, but those users blamed it on the AT, and did not want to report a problem with the website. In the case of partially sighted users, however, users believed that the irresponsiveness was caused by the use of certain graphical elements that "froze" the screen and became irresponsive.

4.4 Accessibility problems encountered by dyslexic users

Dyslexic users encountered a total of 693 problems in this study. Out of those problems, 541 (78.1%) were reported and had their severity rated by users, 71 (10.7%) were mentioned by users but not rated and 78 (11.3%) were identified by the researcher. A total of 43 subcategories had instances of problems encountered by dyslexic users.

4.4.1 Main problems encountered by dyslexic users

From the 43 subcategories of problems encountered by dyslexic users, the 15 most frequent subcategories accounted for 81.5% of all problems. Table 4.7 presents the list of subcategories, followed by the number of instances of problems that occurred and the percentage of the total number of problems accounted by each individual subcategory, and with the median severity rating for each category.

Table 4.7. Median severity rating of the fifteen most frequent subcategories of problems encountered by dyslexic users

Subcategory description	**Instances - N (%)**	**Median Severity Rating**
1. Content not found in pages where expected by users (Content)	112 (16.16)	**3**
2. Navigation elements do not help users find what they are seeking (Navigation)	87 (12.55)	**3**
3. Difficult to scan page for specific item (All media types)	72 (10.39)	2
4. Default presentation of text not adequate (Text)	44 (6.35)	2
5. Expected functionality not present (Controls, forms or functionality)	37 (5.34)	**3**
6. Too much information on page (Content)	34 (4.91)	2
7. Organisation of content is inconsistent with web conventions/common sense (Content)	30 (4.33)	2
8. Users cannot make sense of content (Content)	29 (4.18)	**3**
9. Functionality does not work (as expected) (Controls, forms or functionality)	29 (4.18)	**3**
10. Users cannot understand sequence of interaction (Controls, forms or functionality)	17 (2.45)	**3**
11. No/insufficient feedback to inform that action has had an effect (Controls, forms or functionality)	17 (2.45)	2
12. Link destination not clear (Links)	16 (2.31)	2
13. Language too complicated for perceived target audience (Content)	15 (2.16)	2
14. Irrelevant content before task content (Content)	13 (1.88)	2
15. Users inferred the existence of functionality where there was not one (Controls, forms or functionality)	13 (1.88)	2

Six of the 15 most frequent subcategories of problems encountered by dyslexic users had median severity rating 3 (major). The categories with high frequency and high median severity rating were: "*content not found in pages where expected by users*", "*navigation elements do not help users find what they are seeking*", "*expected functionality not present*", "*users cannot make sense of content*", "*functionality does not work (as expected)*" and "*users cannot understand sequence of interaction*".

4.4.1.1 Content not found in pages where expected by users

Like blind and partially sighted users, dyslexic users also encountered problems when they could not find content in pages where they expected. Those problems were the most frequently encountered by dyslexic users, with 112 instances, accounting for 16.16% of all problems encountered by those users. Problems in this category also had median severity rating 3 – major.

In those problems, users confidently followed a link to a page, but a piece of information that they expected to find there was missing. For example, on a restaurant website, users sought information about the price of a platter for delivery, and found a page with a description of platters. They expected that the page would list all information about platter, including prices, but they could not find any information about prices. Figure 4.13 shows the illustration of the screen, listing two platters, but with no information about prices, as expected by users. The information about prices was available only on a separate PDF document that had to be downloaded.

Figure 4.13 – Example of page listing platters in a menu, but without prices, as expected by users

Many dyslexic users affirmed that making sense of the structure of a website to find their way around is a very important part of their navigation tasks, as it helps them get a sense of direction that is crucial for them to overcome difficulties associated with their dyslexia. Not finding information in places they were confident were the right ones to have what they wanted can be very serious and make them lose the confidence in the mental model they created about the websites.

The main technical cause of problems in this subcategory was connected to the information architecture and organisation of websites. Users expected to find certain pieces of information in parts of the website that were different from that laid-out by the designers of the websites.

4.4.1.2 Navigation elements do not help users find what they are seeking

Encountering navigation elements that do not help users find what they are seeking was also one of the most frequent subcategories of problems encountered by dyslexic users, with 87 instances, which accounted for 12.55% of all problems encountered by those users.

Like blind and partially sighted users, dyslexic users also found problems with navigation elements that were confusing and disorienting, and did not help them find the information they were seeking in their task. In one example of problem in this subcategory, users were seeking tips for driving tests targeted at young people. The navigation offered several options that seemed to be plausible, such as “Motoring Advice” and “Driving School””, but users could not be sure which one to follow. Figure 4.14 shows the navigation bar of The Automobile Association website with the options available.

Figure 4.14 – Example of navigation of The Automobile Association website – users had difficulties finding where to find information about driving tests for young people

Dyslexic users expressed that well-structured navigation and organisation of information websites is fundamental to them. Several dyslexic users stated that websites with information poorly organised made it very hard for them to form a mental model of the website, and that good structuring helps immensely with their processing of information to locate what they are seeking in websites.

4.4.1.3 *Expected functionality not present*

Users reported problems when they expected websites to have a functionality that was not present. Those problems were among the most frequently encountered by dyslexic users, with 37 instances, or 5.34% o f all problems encountered by those users. Those problems also had median severity rating 3 (major).

The most frequent problems reported by users were the lack of search features and features of auto-completion in form fields. In the case of the search features, many users reported that they prefer to use a search when using websites to speed up their navigation. When a search feature was not available, they reported that this prevented them from using an essential strategy to find information that they were most used to.

Users with spelling difficulties reported that they would benefit significantly from having an auto-completion feature in form fields. With this feature, they could type the first few letters of the word they wanted, and the system would provide them with suggestions of words that begin with the letters they typed. For users with more severity spelling difficulties, this can mean that they have to repeat operations where they input content several times in order to search for an item, for example, or that they can have spelling mistakes in input that can compromise the data they provide in a website. Figure 4.15 illustrates an example of an input field with an auto-completion feature that provides suggestions of words that start with the letters pressed by the user.

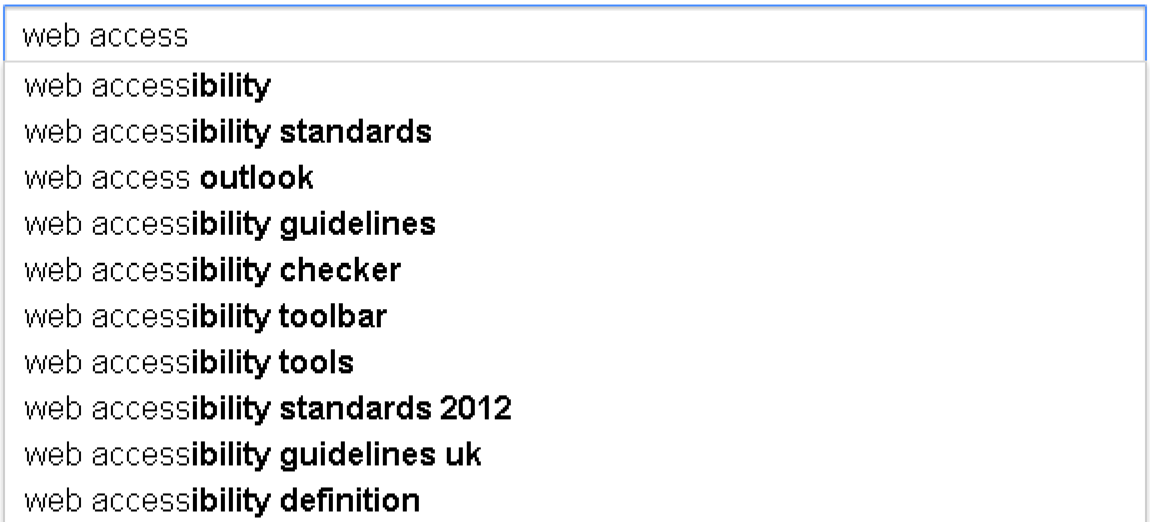


Figure 4.15 – Example of auto-complete feature in input fields that helps dyslexic users with spelling difficulties to complete words

The use of auto-complete features seems to be growing in a number of websites, especially in widely used search engines. Dyslexic users reported that this is a very useful aid for them, as they do not always remember how to spell certain words, and benefit from having the system help complete words.

Search seems to be a key feature to users, as many of them prefer to use it as a first strategy to find content on websites. Especially for dyslexic users who had difficulties forming a model of the structure of the website, having a search feature would greatly help them arrive at the content they want without having the mental load of abstracting the website's organisation.

4.4.1.4 Users cannot make sense of content

Users not being able to make sense of content accounted for 4.18% of the problems with 29 instances, and had high median severity rating, and the subcategory "*English was too complicated for the perceived target audience*" accounted for 2.16% of the problems. Both categories were strongly related to the use of difficult language in the content of websites. The results relates to findings from previous studies that highlight the importance of using clear language to make it easier for dyslexic users to read (Bradford 2005, British Dyslexia Association 2011, Disability Rights Commission 2004, Zarach 2002).

Users often encountered incomplete information or unclear explanations about content that was specific to an unfamiliar domain in a website. On the TicketMaster website, users were trying to buy tickets for an event. They encountered a seat plan with circles with different colours that represented seats. Available seats were orange and unavailable seats were grey. Users tried to click on several different seats using trial and error to figure out which seats were available, since they could not infer this just by the colours that were used.

4.4.1.5 Functionality does not work (as expected)

Problems with malfunctioning features were also among the most frequent problems encountered by dyslexic users, with 29 instances (4.18% of all problems) and median severity rating 3 (major). The most notorious of those problems with dyslexic users was with search features that did not return any results with keywords that users expected would be in the website.

As discussed previously about the importance many dyslexic users place on search features, having ineffective searches can seriously erode users' confidence on the information finding mechanisms made available on a website. If users try to use searches successive times for information they believe are on the website and they are not found, they may become less confident on the website and spend more time using other strategies that are not their preferred option.

The main cause of such problems was that many websites included only a fraction of web pages in the index of the search, not including static pages, for example.

Other problems were related to functionality that simply did not respond to users' actions or behaved in an unexpected manner, which correspond to features that were not properly tested to ensure the correctness of the implementation of features by developers.

4.4.2 Other important features of problems encountered by dyslexic users

4.4.2.1 Highlighting and scanning information

The subcategory "*difficult to scan page for specific item*" was one of the most frequent types of problems, which accounted for 10.4% of the problems encountered by dyslexic users. This category is related to problems when the user encounters difficulties scanning for specific items in a web page, often due to lack of structural or visual aids that would make the content they needed stand out from the rest of the web page. When proper highlighting and structural elements are not found, scanning for information becomes a considerably more difficult task to dyslexic users. This also relates to results from other studies that indicate that appropriate highlighting strategies are very important to help dyslexic users scan for information in websites (Bradford 2005, British Dyslexia Association 2011, Disability Rights Commission 2004, Rello et al. 2012, Zarach 2002).

4.4.2.2 Amount of information and organisation of content

The results from this study also confirmed the importance of avoiding too much information and providing good organisation of content in a web page for dyslexic users, which were pointed by previous studies (Bradford 2005, British Dyslexia Association

2011, Disability Rights Commission 2004, Rello et al. 2012, Zarach 2002). Users being bombarded with too much information in a page accounted for 4.91% of the total number of problems. Illogical organisation of information within a web page accounted for a further 4.33%, and irrelevant information before relevant content accounted for 1.88% of the problems.

4.4.2.3 Problems with presentation of text

Issues with the way text was presented were among the fifteen most frequent types of problems, with 6.83% of all problems encountered by dyslexic users. The main problems encountered were in line with existing layout guidelines for dyslexic users (Bradford 2005, British Dyslexia Association 2011, Disability Rights Commission 2004, Rello et al. 2012, Santana et al. 2012, Zarach 2002): use of italics, inadequate spacing between lines and paragraphs, small font size, inappropriate font style, text in columns and inappropriate colour contrast.

This subcategory had median severity 2 (minor). However, users encountered a substantial number of catastrophic and major problems as well. This reflects how different users can be affected by problems with the presentation of text depending on the severity and type of their dyslexia. Many users reported the issue with colour contrast as being minor, for example, but there were extreme cases in which users reported developing headaches and had to stop completely a task after reading a long text with black writing on white background.

4.4.2.4 Customising colour contrast

Problems with reading black writing on white background were reported by seven of the 13 dyslexic users that took part in this study. However, none of the participants reported to use any specific assistive technology to help them change colours to use websites. It is also noteworthy that none of the users used the web browser's settings to change this either. Although participants who use such technologies might not have been included in the present study, the results from the sample showed that there can be a substantial number of users who need means to change colour schemes other than specialised assistive technologies.

4.5 Distinct website problems encountered by different user groups

In order to be able to compare problems encountered by different users (within and between user groups), most analyses in this section are based on *distinct website problems*, which correspond to a problem in a website that may have occurred in more than one instance to different users, as described in Section 3.6.4. Distinct website problems were related to problems connected to the same interface component in a website, and which caused problems of the same nature to users in different instances.

The 3,012 *instances* of user problems were mapped onto a total of 1,513 *distinct website problems*. Distinct website problems had between 1 and 30 instances each. Table 4.8 shows the number of distinct website problems that had were encountered by blind, partially sighted and dyslexic users. The table contains the number of distinct website problems and the percentage of problems that were encountered by each group in relation to all distinct website problems encountered in the study.

Table 4.8. Number of distinct website problems per user group

User group	Number of Distinct Website Problems – N (%)
Blind	771 (51%)
Partially sighted	649 (42.9%)
Dyslexic	412 (27.2%)

The Venn diagram in Figure 4.16 shows the number and percentage of distinct website problems that were encountered by each user group and problems that were common to two or more groups. More than 83% of the distinct website problems were encountered exclusively by one single user group. Another 17.1% of the distinct website problems were encountered by users from two or more user groups

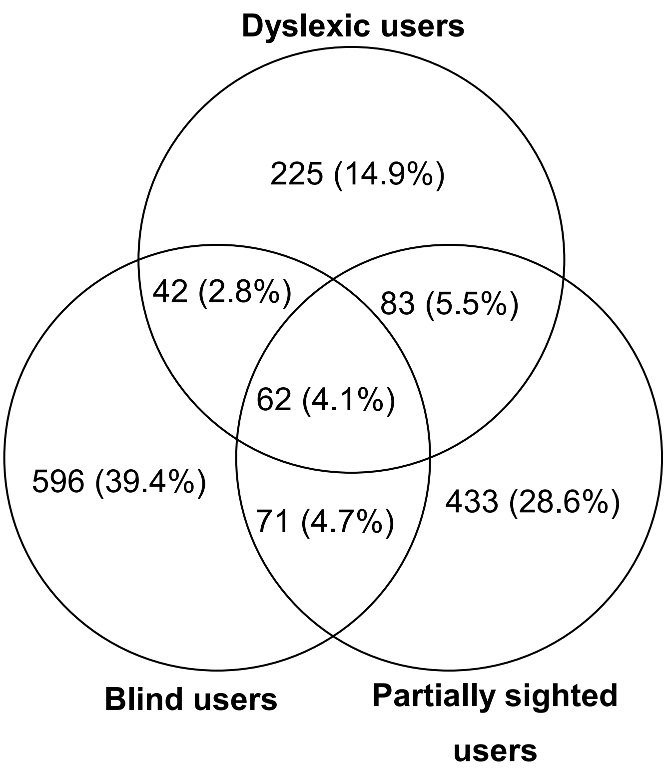


Figure 4.16 – Venn diagram with distinct website problems encountered by different user groups

4.5.1 Distinct website problems encountered by all user groups

Table 4.9 presents a list of subcategories of problems and the number of distinct problems in each subcategory that were encountered by users in all user groups. The subcategories "*multimedia starting automatically is irritating"* and "*inadequate alternative to functionality*" only had one distinct website problem each, which means they only occurred in one website and only once. However, it is worth noting that instances of problems in these subcategories occurred to users in all user groups. Problems related to "*meaning in content is lost or modified due to transformation"* had 50% of the distinct problems in the subcategories encountered by all user groups.

Table 4.9. Subcategories with distinct website problems encountered by all user groups

Subcategory description	**Distinct Problems encountered by all groups – N (% of subcategory)**
Multimedia starting automatically is irritating	1 (100)
Inadequate alternative to functionality	1 (100)
Meaning in content is lost or modified due to transformation	2 (50)
Users cannot understand sequence of interaction	2 (20)
Navigation elements do not help users find what they are seeking	16 (18.2)
Content not found in pages where expected by users	13 (14.4)
Navigation elements not understandable (e.g. using jargon of difficult language)	1 (14.3)
Irrelevant content before task content	4 (10.3)
No cue regarding specific input requirements	1 (9.1)
Users inferred the existence of functionality where there was not one	1 (9.1)
Language too complicated for perceived target audience	1 (7.7)
No/insufficient feedback that users have not performed a required action	1 (6.7)
Link grouping poor	1 (6.3)
No obvious way to return to homepage	1 (5.9)
System too slow	1 (5.6)
Functionality does not work (as expected)	3 (5.6)
No/insufficient feedback to inform that action has had an effect	2 (5)
Information architecture too complex	1 (4.3)
Expected functionality not present	2 (3.8)
Difficult to scan page for specific item	2 (2.8)
Too much information on page	1 (2.8)
Link destination not clear	2 (2.1)

4.5.2 Distinct website problems encountered only by blind users

A total of 49 subcategories had the majority of their distinct website problems encountered by blind users only, ranging from 24.7% to 100% of the problems in the subcategory encountered by blind users only. It is noteworthy that of those 49 subcategories, 16 of them had 100% of all their distinct website problems encountered exclusively by blind users.

Table 4.10 presents a list of the 16 subcategories with 100% of distinct problems encountered only by blind users. It is noteworthy that problems in several subcategories related to headings were only encountered by blind users: "*no headings*", "*heading structure violated*", "*heading content not available*", "*headings do not give overview*", "*no heading when one is needed*" and "*heading not perceived as being a heading*". Two subcategories related to tables also had problems that were only encountered by blind users: "*table cell not associated with headers*" and "*no heading to identify table columns/rows*". Other subcategories related to images and links also had 100% of their distinct problems only encountered by blind users.

Table 4.10. Selection of 16 subcategories of problems with 100% of distinct website problems encountered only by blind users

Subcategory description	**Distinct Problems encountered by blind users only – N (% of subcategory)**
No heading when one is needed (Headings)	20 (100)
Inadequate alternative to image (Images)	19 (100)
No headings (Headings)	19 (100)
Heading not perceived as being a heading (Headings)	9 (100)
No alternative to image (Images)	8 (100)
Heading structure violated (Headings)	8 (100)
Too many links (Links)	7 (100)
Headings do not give overview (Headings)	5 (100)

Subcategory description	Distinct Problems encountered by blind users only – N (% of subcategory)
Link destination not present (Links)	5 (100)
Repeated links (Links)	5 (100)
Users inferred the existence of functionality where there was not one (Controls, forms or functionality)	4 (100)
Users cannot associate table cell with headers (Tables)	4 (100)
No heading to identify table columns/rows (Tables)	2 (100)
System executes action unexpectedly (Controls, forms or functionality)	2 (100)
Too many headings (Headings)	1 (100)
Heading content not available (Headings)	1 (100)

4.5.3 Distinct website problems encountered only by partially sighted users

A total of 18 subcategories had the majority of their distinct website problems encountered by partially sighted users only, ranging from 40% to 100% of the problems in the subcategory being encountered by partially sighted users only. Table 4.11 presents the list of these subcategories.

As can be seen in Table 4.11, two subcategories had 100% of their unique website problems encountered exclusively by partially sighted users: "*assistive technology becomes irresponsive with particular graphic elements*" and "*link target area not operable*". Blind users did not report problems when their assistive technology became irresponsive, as they normally blamed it on the AT and not on the graphical components on websites.

It is also noteworthy that a substantial percentage of problems related to inadequate presentation of elements were encountered exclusively by partially sighted users: "*default presentation of image not adequate*" (96.6%), "*default presentation of control or* *form element not adequate*” (90.3%), “*default presentation of text not adequate*” (75.3%) and “*default presentation of table not adequate*” (66.7%).

Table 4.11. Subcategories of problems with majority of distinct website problems encountered only by partially sighted users

Subcategory description	**Distinct Problems encountered by partially sighted users only – N (% of subcategory)**
Default presentation of audio, video or multimedia not adequate (Audio, video or multimedia)	5 (100)
Assistive technology becomes irresponsive with particular graphic elements (Underlying system characteristics)	4 (100)
Graphic or multimedia not compatible with assistive technology (Underlying system characteristics)	2 (100)
Link target area not operable (Links)	1 (100)
Default presentation of image not adequate (Images)	29 (96.7)
Default presentation of control or form element not adequate (Controls, forms or functionality)	28 (90.3)
Moving multimedia content is distracting (Audio, video or multimedia)	4 (80)
Navigation moves unexpectedly on the screen (Navigation)	4 (80)
Default presentation of text not adequate (Text)	113 (75.3)
Page to page navigation does not go to top of page (Navigation)	2 (66.7)
Default presentation of table not adequate (Tables)	2 (66.7)
Inability to change presentation of audio, video or multimedia (Audio, video or multimedia)	10 (62.5)
Inability to change presentation of text (Text	15 (55.6)
Lack of clearly defined navigation structure (Navigation)	3 (50)
Link grouping poor (Links)	7 (43.8)
Information architecture too complex (e.g. too many steps to find pages) (Information architecture)	10 (43.5)
No alternative to functionality (Controls, forms or functionality)	1 (33.3)

4.5.4 Distinct website problems encountered only by dyslexic users

A total of 6 subcategories had the majority of their distinct website problems encountered by blind users only, ranging from 27.8% to 63.6% of the problems in the subcategory encountered by blind users only. Table 4.12 presents a list of these subcategories.

Table 4.12. Subcategories of problems with majority of distinct website problems encountered only by dyslexic users

Subcategory description	**Distinct Problems encountered by dyslexic users only – N (% of subcategory)**
Navigation bar not salient (Navigation)	7 (63.6)
Language too complicated for perceived target audience (Content)	5 (38.5)
Expected functionality not present (Controls, forms or functionality)	20 (38.5)
Difficult to scan page for specific item (All media types)	27 (38)
Users inferred the existence of functionality where there was not one (Controls, forms or functionality)	4 (36.4)
Too much information on page (Content)	10 (27.8)

As can be observed in Table 4.12, the subcategory "*navigation bar not salient*" had the highest percentage of distinct problems encountered exclusively by dyslexic users (63.6%), followed by "*difficult to scan page for specific item*" (38). Issues with difficult language, expected functionality not present, inference of existence of functionality and too much information on page also had a substantial percentage of problems reported exclusively by dyslexic users. Most problems related to "*expected functionality not present"* referred to the lack of auto-complete features that would help users with spelling difficulties.

4.6 Severity of problems encountered by different user groups

Two analyses were performed on the agreement level of severity ratings within the same user group and between user groups for problems in common. Section 4.6.1 presents the analysis of agreement between users from the same user group, and Section 4.6.2 presents an analysis of the difference between severity ratings from users from different user groups.

4.6.1 Agreement between severity ratings by users from the same group

The analysis of agreement in the severity ratings within user groups was performed for distinct website problems that were encountered in at least two instances by different users within a same user group. The agreement level was performed using a method defined by Petrie and Kheir (2007), based on the range of severity ratings assigned to a given problem. Ratings for one problem can be in “total agreement” if all severity ratings are the same or can have: a “1 difference” (1-diff) if all severity ratings differ by no more than 1 (for example, with two ratings 2 – minor and one rating 3 – major), a “2 difference”, if severity ratings differ by no more than 2 (for example, with one rating 1 - cosmetic, one 2 – minor and one 3 – major) or a “3 difference” (for example, with one rating 1 – cosmetic and one 4 – catastrophe).

Figure 4.17 shows a graph with the percentage of problems encountered by two or more blind users in each level of agreement, based on 149 distinct problems. The category “2 difference” had the majority of problems, with 33.6%. The sum of problems in which severity ratings were in total agreement or had a “1 difference” corresponded to 53.7% of the problems.

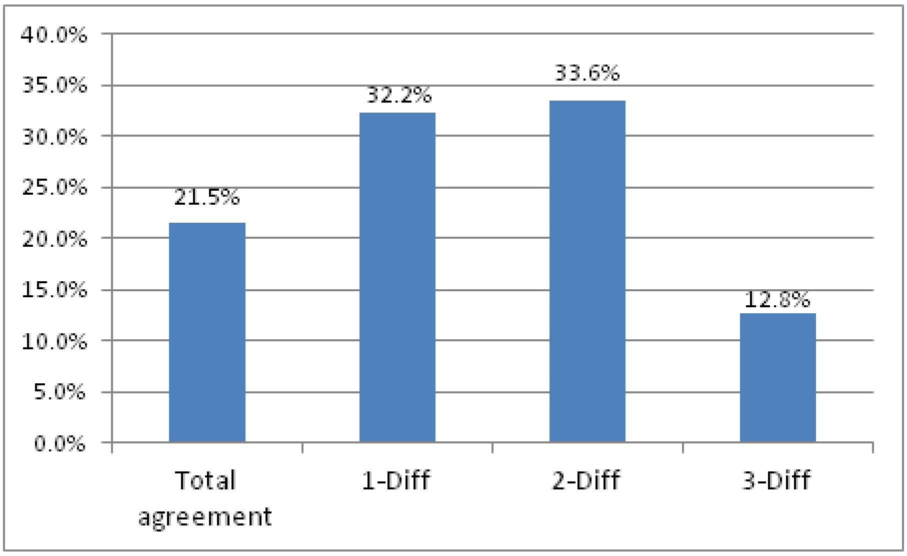


Figure 4.17 – Agreement between blind users in problem severity ratings

Figure 4.18 shows a graph with the percentage of problems encountered by two or more partially sighted users in each level of agreement, based on 123 distinct problems. The category “1 difference” had the majority of problems, with 40.7%. The sum of problems in which severity ratings were in total agreement or had a “1 difference” corresponded to 73.2% of the problems.

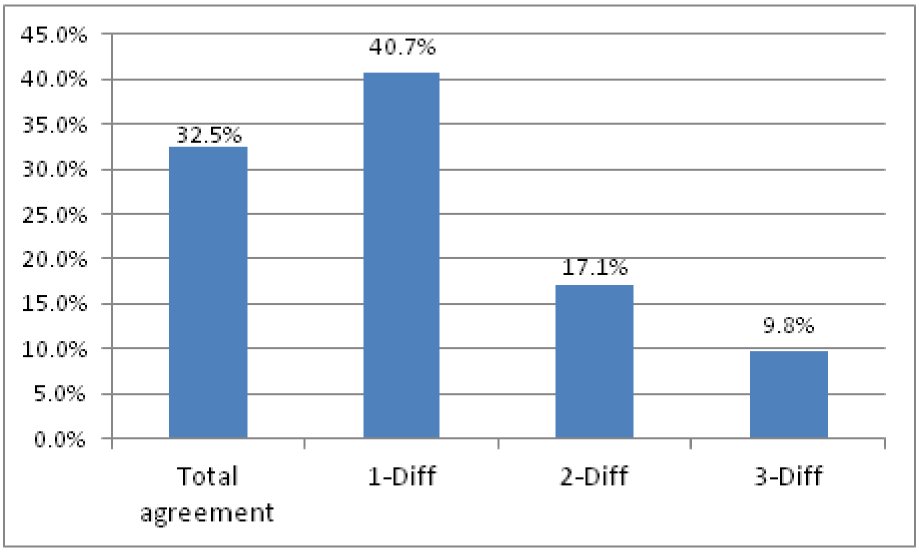


Figure 4.18 – Agreement between partially sighted users in problem severity ratings

Figure 4.19 shows a graph with the percentage of problems encountered by two or more dyslexic users in each level of agreement, based on 97 distinct problems. The category “1 difference” had the majority of problems, with 49.5%. The sum of problems in which severity ratings were in total agreement or had a “1 difference” corresponded to 72.2% of the problems.

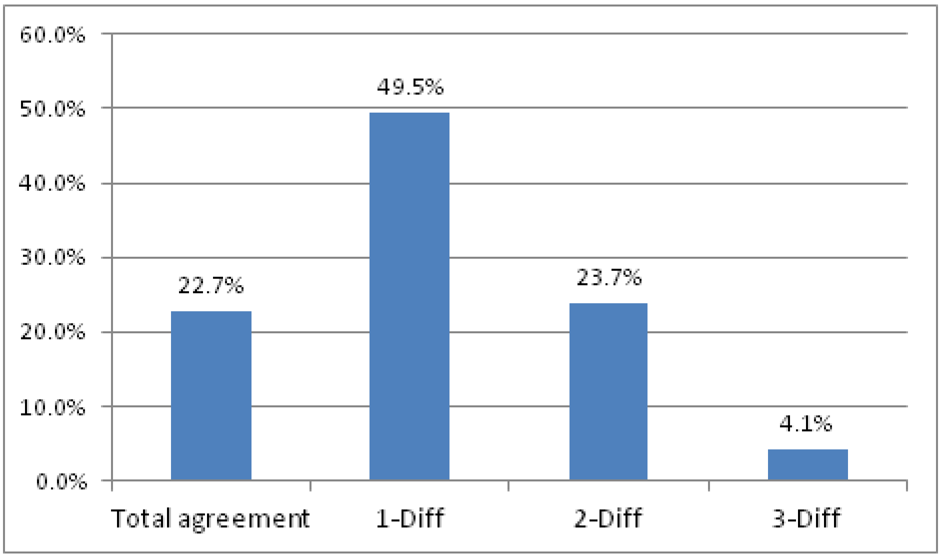


Figure 4.19 – Agreement between dyslexic users in problem severity ratings

Partially sighted and dyslexic users had a higher level of agreement in the severity ratings than blind users.

4.6.2 Difference between severity ratings of problems in common for different user groups

Analyses were performed on the degree of agreement in the severity ratings from different user groups that encountered common distinct website problems, considering those encountered in common by blind and partially sighted, blind and dyslexic and partially sighted and dyslexic users.

The first analysis compared the mean severity ratings of 71 distinct problems encountered by both blind and partially sighted users. The mean severity rating of those problems from blind users was 2.77, with SD 0.78, and partially sighted users had mean 2.98 and SD 0.68. A Related-Samples Wilcoxon Signed Rank test showed a significant difference between ratings of problems in common encountered by blind and partially sighted users ($W_+ = 1.994$, df = 1, $p < 0.05$).

The second analysis compared the mean severity ratings of 42 distinct problems encountered by both blind and dyslexic users. The mean severity rating of those problems from blind users was 2.40, with SD 0.82, and dyslexic users had mean 2.53 and SD 0.76. A Related-Samples Wilcoxon Signed Rank test showed no significant difference between ratings of problems in common encountered by blind and partially sighted users ($W_+ = -0.620$, df = 1, $p = 0.535$).

The third analysis compared the mean severity ratings of 83 distinct problems encountered by both partially sighted and dyslexic users. The mean severity rating of those problems from partially sighted users was 2.85, with SD 0.67, and dyslexic users had mean 2.34 and SD 0.73. A Related-Samples Wilcoxon Signed Rank test showed a significant difference between ratings of problems in common encountered by partially sighted and dyslexic users ($W_+ = -4.052$, $df = 1$, $p < 0.05$).

4.7 Summary of the chapter

This chapter presented the main results related to the primary research question proposed in this research: "*What are the main characteristics of accessibility problems encountered by print-disabled users when attempting to use websites?*" and related sub-questions.

The results presented showed measures of task success, task difficulty and problems found by users. Evidence in the study revealed that print-disabled users still encounter a substantial number of accessibility problems when attempting to perform tasks on websites, and that blind and partially sighted users were particularly affected and could not succeed in a substantial amount of tasks.

The chapter also presented a characterisation of the main types of problems encountered by blind, partially sighted and dyslexic users. Analyses covered the frequency and severity of the main types of problems encountered by those groups. In order to characterise those problems, analyses were performed to identify the most frequent problems, the most severe problems and the severity of the most frequent problems.

Analyses of distinct website problems (that may have had more than one instance) were also performed, in order to perform comparisons between the severity ratings assigned within and between user groups. The analyses showed that partially sighted users had higher severity ratings than blind and dyslexic users.

Other analyses aimed at identifying problems that were encountered in common by different user groups and problems that are particular to individual user groups. They showed that although some problems were common to different user groups, there was a substantial amount of problems that were only encountered by certain users, reinforcing the need to diversify the sample of users that take part in user evaluation to cover as many problems as possible.

The next chapter builds up on the results presented in the current chapter, and aims to answer a secondary research question. It investigates the relationship between the accessibility problems encountered by print-disabled users and technical accessibility guidelines in WCAG 1.0 and 2.0.

Chapter 5. The relationship between problems encountered by print-disabled users and technical web accessibility guidelines

This chapter presents the results and discussion of the findings related to the secondary research question proposed in this research: "*What is the relationship between user-based measures of accessibility of websites and measures of technical web accessibility based on the guidelines defined in the Web Content Accessibility Guidelines 1.0 and 2.0?*". The chapter provides different analyses comparing problems encountered by print-disabled users with different aspects of technical the Web Content Accessibility Guidelines (WCAG) 1.0 and 2.0.

Section 5.1 presents the results of comparisons between numbers of user problems and levels of conformance to WCAG. Section 5.2 shows the comparison between user problems and instances of violations and numbers of different WCAG 1.0 checkpoints (CPs) or WCAG 2.0 success criteria (SCs) violated. Section 5.3 presents the analysis of the coverage of problems encountered by different user groups by WCAG 1.0 and WCAG 2.0. Section 5.4 describes the main types of user problems not covered by WCAG 1.0 and WCAG 2.0. Section 5.5 shows the main types of user problems that were covered by WCAG 1.0 or WCAG 2.0, and users still encountered problems despite related checkpoints or success criteria having been successfully implemented by websites. Section 5.6 presents an analysis of the relationship between the priority of WCAG 1.0 CPs/WCAG 2.0 SCs and the severity ratings of problems assigned by users or by researchers. Finally, Section 5.7 presents a summary of the chapter.

5.1 User problems and WCAG conformance

The first analysis comparing problems encountered by print-disabled users and WCAG considered the levels of conformance of the home pages of websites. Results from previous studies (Disability Rights Commission 2004) have established a very high correlation of the WCAG 1.0 conformance of the home page of a website with the

WCAG 1.0 conformance of other pages of that same website, as described in Section 2.5.1. For this reason, the comparisons in this and in the next section are based on the audits of the home pages of websites.

The comparison considered the mean number of user problems encountered per website per user. Table 5.1 presents a list of websites and information about their level of conformance to WCAG 1.0 and WCAG 2.0. For each website, the level of conformance (column Conf. with values A, AA, AAA and Fail for failure) is shown, along with the number of CPs/SCs violated and the number of instances of violations (column *Inst.*). The table is ordered in descendent order of WCAG 1.0 conformance levels.

The *instances of violations* of CPs/SCs (*Inst.*) is the number of times each CPs/SCs is violated. For example, a website has two different CPs/SCs violated: there are 10 images without alternative text (CP 1.1/SC 1.1.1) and 22 missing headings (CP 5.1/SC 2.4.6). The instances of violations is 32.

Table 5.1 shows that only 1 website with WCAG 2.0 Level AAA and 1 website with WCAG 2.0 Level AA could be found. In addition, a number of sites, which conformed to WCAG 1.0 at A or AA, failed WCAG 2.0 conformance. When examining the ways in which websites failed to conform, it was noted that 3 of the websites (York, NHSNSS and The AA) failed one SC a single time, specifically SC 3.3.1 (“error identification”), and otherwise conformed to Level A of WCAG 2.0. Therefore, in some analyses, which will be noted, these websites are classified as Level A conformant websites.

5.1.1 Problems encountered by blind users and WCAG conformance

For blind users, the mean number of problem instances per website across all websites was 9.22, with SD 5.31. The website Ticketmaster (www.ticketmaster.co.uk) had mean number of problems per user above two times the standard deviation across all websites and was considered an outlier. This website was omitted from all analyses considering the mean number of problems per website per user for blind users.

Table 5.1. List of websites evaluated in the study, with conformance levels to WCAG 1.0 and WCAG 2.0, number of different CPs/SCs violated and instances of violations, and instances of user problems grouped by severity levels

Website	WCAG 1.0			WCAG 2.0			Mean number of problems per website per user		
	Conf.	CPs violated	Inst.	Conf.	SCs violated	Inst.	Blind	Partially sighted	Dyslexic
www.lflegal.com	AA	2	5	AAA	0	0	4.6	6.5	4.9
www.green-beast.com	AA	3	23	AA	3	9	4.9	5.3	3.2
www.york.gov.uk	AA	4	16	Fail	5	7	7.0	10.6	6.3
www.nhsnss.org	AA	6	30	Fail	9	31	9.2	7.3	4.6
www.copac.ac.uk	A	8	21	A	2	6	5.7	4.3	2.9
www.theaa.com	A	9	68	Fail	9	58	5.4	8.9	3.7
www.dh.gov.uk	A	9	91	A	6	31	9.4	6.4	6.9
www.digizen.org.uk	A	9	80	Fail	12	46	5.4	7.9	1.3
www.jisc.ac.uk	A	12	58	Fail	13	216	5.1	6.3	4.6
www.royalmail.com	A	15	50	Fail	7	103	5.4	7.4	5.0
www.pret.co.uk	A	16	184	Fail	21	141	5.4	5.3	1.8
www.tuc.org.uk	A	23	146	Fail	17	97	10.0	10.0	5.7
www.britishmuseum.org	Fail	8	130	Fail	8	86	14.2	10.5	5.7
www.nhsdirect.nhs.uk	Fail	10	30	Fail	20	163	7.3	5.0	4.4
www.ford.co.uk	Fail	27	124	Fail	33	244	16.9	9.6	5.2
www.ticketmaster.co.uk	Fail	29	757	Fail	35	1118	19.8	11.0	3.1

Figure 5.1 shows the mean number of problems per website per blind user grouped and its variation grouped by the level of conformance to WCAG 1.0 of their home pages. A one-way ANOVA showed a significant difference in the number of problems per website in different WCAG 1.0 conformance levels ($F = 12.35$, $df = 2, 13$, $p < 0.001$). Follow-up Tukey HSD post-hoc tests showed a significant difference between non-conformant and Level A ($p < 0.001$) and non-conformant and level AA websites ($p < 0.005$).

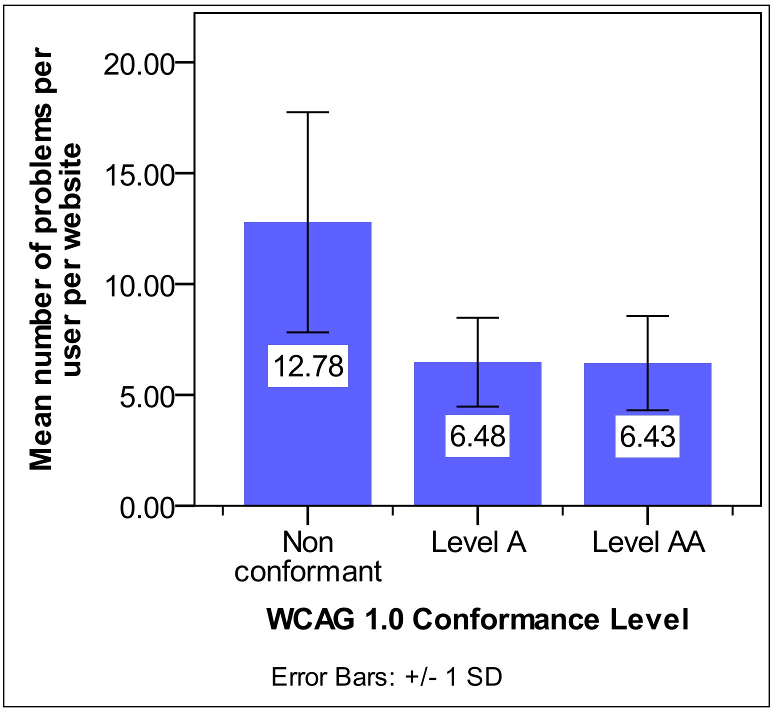


Figure 5.1 – Blind users - mean number of instances of user problems per website per user grouped by WCAG 1.0 conformance levels

Figure 5.2 shows the mean number of problems per website per blind user grouped and its variation grouped by the level of conformance to WCAG 2.0.of their home pages. It was not possible to make a comparison between the individual levels of conformance as there were so few websites that conformed to Level AA or Level AAA (see Table 5.1).

A comparison was made between non-conformant websites and websites that were conformant to WCAG 2.0 at any level (levels A, AA and AAA). Across all websites that were not conformant to WCAG 2.0, the mean number of problems per user per website was 9.42, with SD 4.17, whilst the mean for websites conformant at any level was 5.79, with SD 1.61, and for websites at level A only, the mean was 6.2, with SD 1.8.

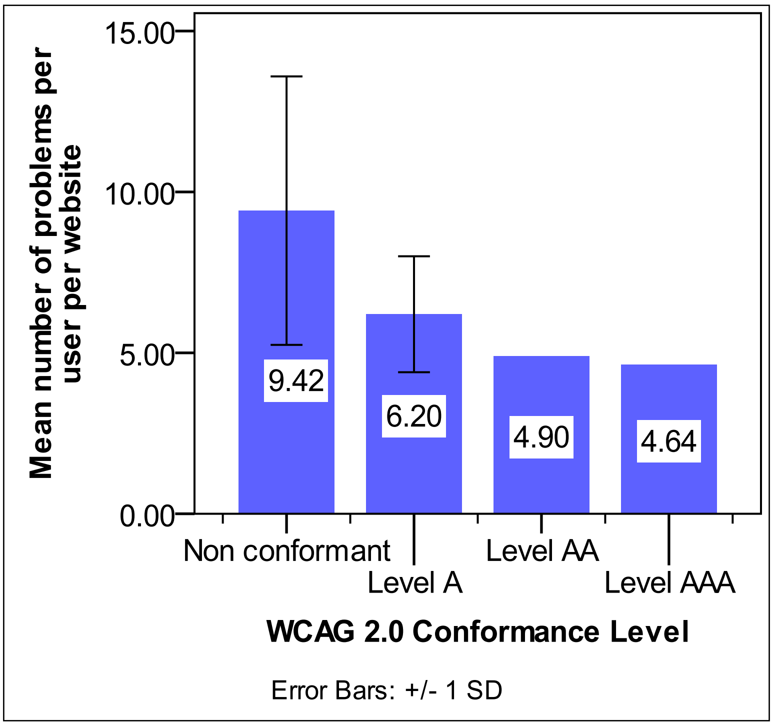


Figure 5.2 – Blind users - mean number of instances of user problems per website per user grouped by WCAG 2.0 conformance levels

For the comparisons with WCAG 2.0, two separate tests were performed. The first test found no significant difference between the mean number of problems in non-conformant websites and Level A conformant websites (F = 1.107, df = 1,12, p = 0.544). The second test also found no significant difference between the mean number of problems in non-conformant and conformant websites at all levels (F = 1.258, df = 1,14, p = 0.282). In these analyses the three websites which failed SC 3.3.1 on one occasion only were classified as Level A, as described in Section 5.1.

5.1.2 Problems encountered by partially sighted users and WCAG conformance

For partially sighted users, the mean number of user problems per website across all websites was 8.09, with SD 3.18.

Figure 5.3 shows the mean number of problem instances per website per partially sighted user and its variation grouped by the level of conformance to WCAG 1.0.of their home pages. A one-way ANOVA showed no significant difference between the mean number of problems in non-conformant and websites conforming to WCAG 1.0 at levels A and AA (F = 1.096, df = 1,14, p = 0.352).

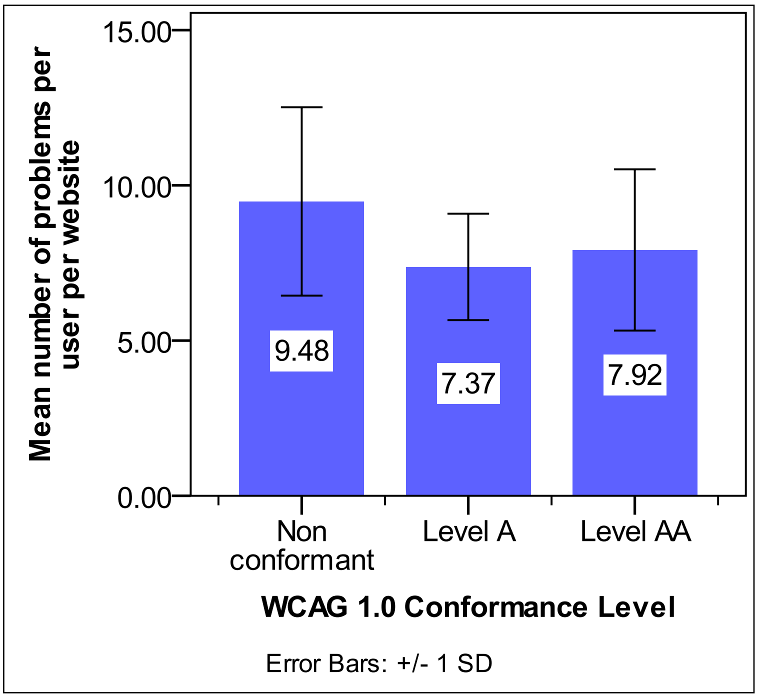


Figure 5.3 – Partially sighted users - mean number of instances of user problems per website per user grouped by WCAG 1.0 conformance levels

Figure 5.4 shows the mean number of problems per website per partially sighted user and its variation grouped by the level of conformance to WCAG 2.0.of their home pages. It was not possible to make a comparison between the individual levels of conformance as there were so few websites that conformed to Level AA or Level AAA (see Table 5.1).

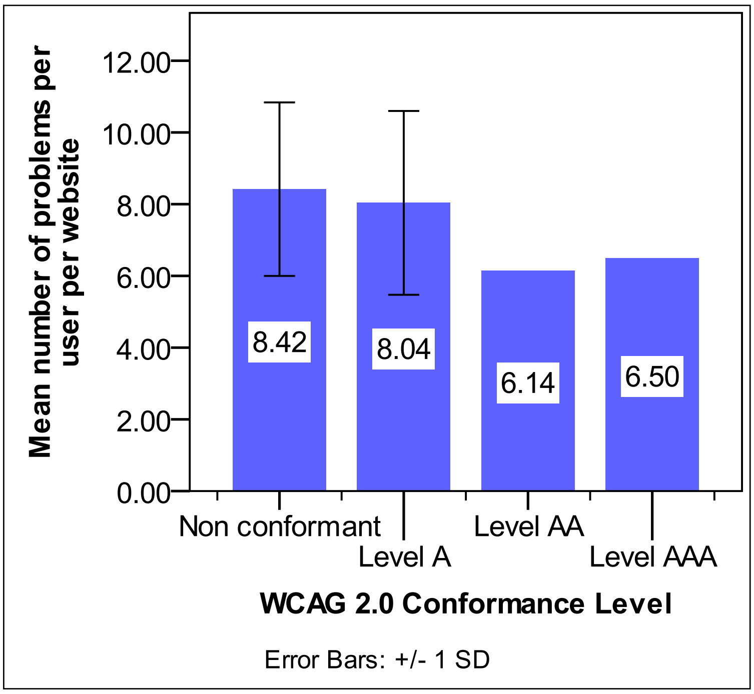


Figure 5.4 – Partially sighted users - mean number of instances of user problems per website per user grouped by WCAG 2.0 conformance levels

A comparison was made bewteen non-conformant websites and websites that were conformant to WCAG 2.0 at any level (levels A, AA and AAA). Across all websites that were not conformant to WCAG 2.0, the mean number of problems per user per website was 8.42, with SD 2.42, whilst the mean for websites conformant at any level was 7.55, with SD 2.26, and for websites at level A only, the mean was 8.04, with SD 2.56.

For this analysis, two separate tests were performed. The first test found no significant difference between the mean number of problems in non-conformant websites and websites at all conformance levels ($F = 0.053$, $df = 1,14$, $p = 0.473$). The second test also found no significant difference between the mean number of problems in non-conformant websites and Level A conformant websites ($F=0.077$, $df=1,12$, $p=0.786$). In these analyses the three websites which failed SC 3.3.1 on one occasion only were classified as Level A, as described in Section 5.1.

5.1.3 Problems encountered by dyslexic users and WCAG conformance

For dyslexic users, the mean number of user problems per website across all websites was 4.64, with SD 1.33.

Figure 5.5 shows the mean number of problems per website per dyslexic user and its variation grouped by the level of conformance to WCAG 1.0.of their home pages. A one-way ANOVA showed no significant difference between the mean number of problems in non-conformant and websites at levels A and AA ($F = 0.362$, $df = 2,13$, $p = 0.703$).

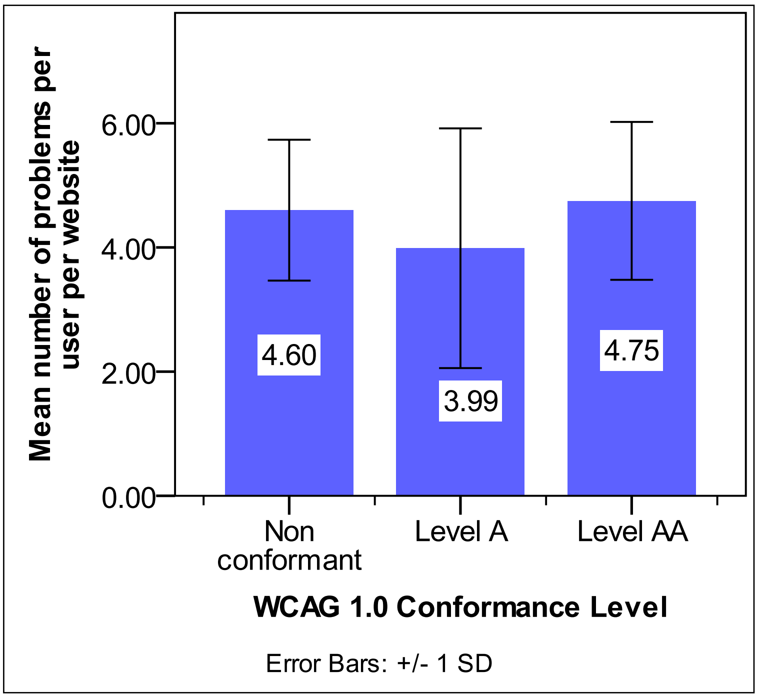


Figure 5.5 – Dyslexic users - mean number of instances of user problems per website per user grouped by WCAG 1.0 conformance levels

Figure 5.6 shows the mean number of problems per website per dyslexic user and its variation grouped by the level of conformance to WCAG 2.0.of their home pages. It was not possible to make a comparison between the individual levels of conformance as there were so few websites that conformed to Level AA or Level AAA (see Table 5.1).

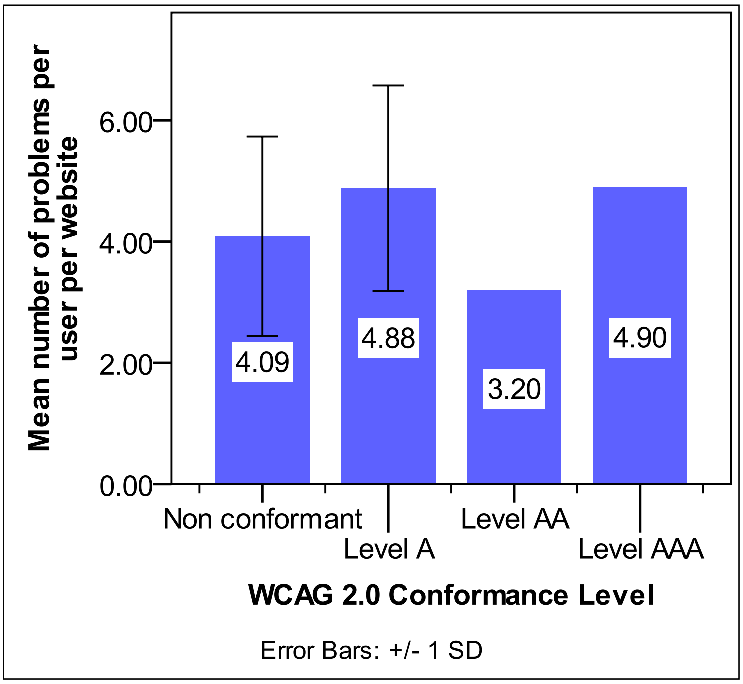


Figure 5.6 – Dyslexic users - mean number of instances of user problems per website per user grouped by WCAG 2.0 conformance levels

A comparison was made bewteen non-conformant websites and websites that were conformant to WCAG 2.0 at any level (levels A, AA and AAA). Across all websites that were not conformant to WCAG 2.0, the mean number of problems per user per website was 4.09, with SD 1.64, whilst the mean for websites conformant at any level was 4.64, with SD 1.52, and for websites at level A only, the mean was 4.88, with SD 1.69.

For this analysis, two separate tests were performed. The first test found no significant difference between the mean number of problems in non-conformant and websites in all levels of conformance ($F = 0.476$, $df = 1,14$, $p = 0.501$). The second test also found no significant difference between the mean number of problems in non-conformant and in level A websites ($F = 0.729$, $df = 1,12$, $p = 0.410$). In these analyses the three websites which failed SC 3.3.1 on one occasion only were classified as Level A, as described in Section 5.1.

5.2 User problems and WCAG checkpoints/success criteria violated

The analysis of between the number of user problems encountered on websites and WCAG CPs/SCs violated was performed using two different measures: the number of *instances of violations* of CPs/SCs and the number of *different CPs/SCs violated*. The following sections present the results of the analyses of those measures comparing to the number of problems encountered by blind, partially sighted, and dyslexic users, respectively.

5.2.1 Problems encountered by blind users and WCAG checkpoints/success criteria violated

For blind users, the analysis between the mean number of problems per website per user and the *instances of violations* of WCAG 1.0 CPs found no significant correlation ($r = 0.474$, $N = 15$, $p = 0.075$). The same was true for the correlation between the number of instances of violations of WCAG 2.0 SCs and the mean number of problems per website per user ($r = 0.405$, $N=15$, $p = 0.134$).

There was, however, a significant correlation between the number of *different WCAG 1.0 CPs violated* and the mean number of problems per website per user ($r = 0.534$, $N=15$, $p < 0.04$). A significant correlation was also found between the number of *different WCAG 2.0 SCs violated* and the mean number of problems per website per user ($r = 0.544$, $N = 15$, $p < 0.036$).

5.2.2 Problems encountered by partially sighted users and WCAG checkpoints/success criteria violated

For partially sighted users, the analysis between the mean number of problems per website per user and the *instances of violations* of WCAG 1.0 CPs found no significant correlation ($r = 0.489$, $N=16$, $p = 0.054$). The same was true for the correlation between the number of *instances of violations* of WCAG 2.0 SCs and the mean number of problems per website per user ($r = 0.420$, $N=16$, $p = 0.106$).

Unlike for blind users, there was no significant correlation between the number of *different WCAG 1.0 CPs violated* and the mean number of problems per website per partially sighted user ($r = 0.453$, $N=16$, $p = 0.078$). No significant correlation was found between the number of *different WCAG 2.0 SCs violated* and the mean number of problems per website per user ($r = 0.400$, $N=16$, $p = 0.125$).

5.2.3 Problems encountered by dyslexic users and WCAG checkpoints/success criteria violated

For dyslexic users, the analysis between the mean number of problems per website per user and the *instances of violations* of WCAG 1.0 CPs found no significant correlation ($r = -0.220$, $df = 15$, $p = 0.413$). The same was true for the correlation between the number of *instances of violations* of WCAG 2.0 SCs and the mean number of problems per website per user ($r = -0.194$, $N=16$, $p = 0.472$).

Unlike for blind users, for dyslexic users there was no significant correlation between the number of *different WCAG 1.0 CPs violated* and the mean number of problems per website per user ($r = -0.054$, $N=16$, $p = 0.842$). No significant correlation was found between the number of *different WCAG 2.0 SCs violated* and the mean number of problems per website per partially sighted user ($r = -0.183$, $N=16$, $p = 0.498$).

5.3 User problems and coverage by WCAG checkpoints/success criteria

In order to establish whether WCAG 1.0 and WCAG 2.0 covered the problems encountered by print-disabled users, an analysis of each user problem was performed in order to match problems with CPs/SCs that could be relevant to each user problem, according to the method described in Section 3.6.5.

For each user problem, it was established whether there were relevant CPs/SCs in WCAG 1.0 and 2.0. If was also analysed whether relevant CPs/SCs clearly addressed the issue related to a given user problem. In this case, the CPs/SCs would be classified as *directly relevant* to the user problem, and would be said to *cover* the problem.

The following sections present the coverage analysis of problems encountered by blind, partially sighted and dyslexic users, respectively.

5.3.1 Problems encountered by blind users and coverage by WCAG checkpoints/success criteria

For WCAG 1.0, **Error! Reference source not found.** shows the breakdown of problems encountered by blind users into categories of relevance of CPs and whether those CPs have been implemented on the website where the problem was encountered. The total percentage of problems encountered by blind users that were covered by CPs was 42.3% (the sum of bars 2 and 3 in **Error! Reference source not found.**) and only a small percentage of those were implemented by developers (3.3% of all user problems, bar 3 in **Error! Reference source not found.**, or 7.7% of all user problems covered by WCAG 1.0). This means that of the problems encountered by users on websites, well over half (57%) were not covered by WCAG 1.0.

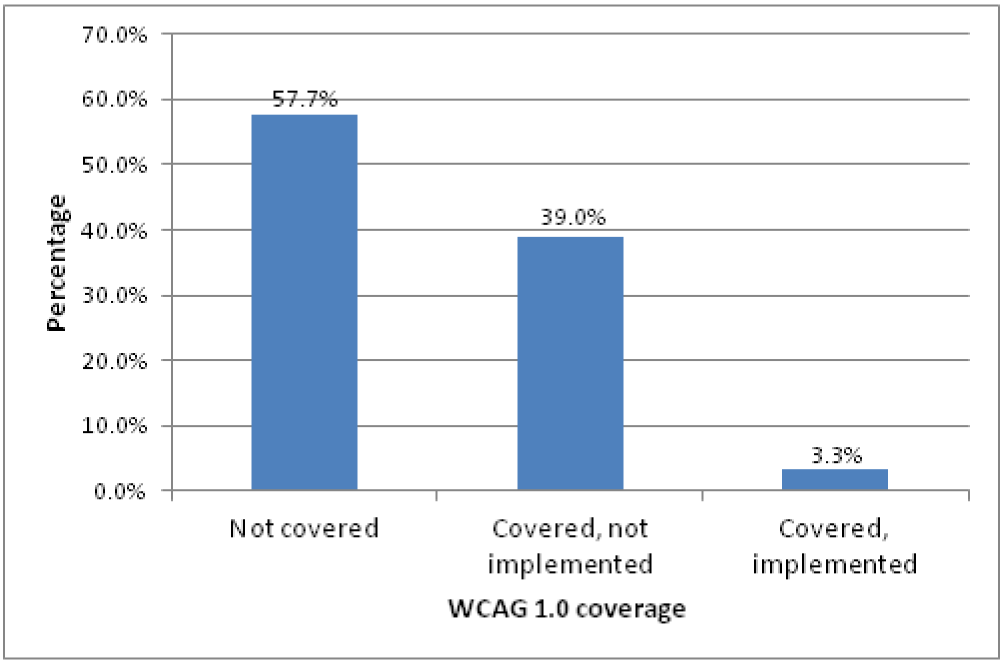


Figure 5.7 –Categories of problems encountered by blind users divided by relevance of WCAG 1.0 CPs and implementation.

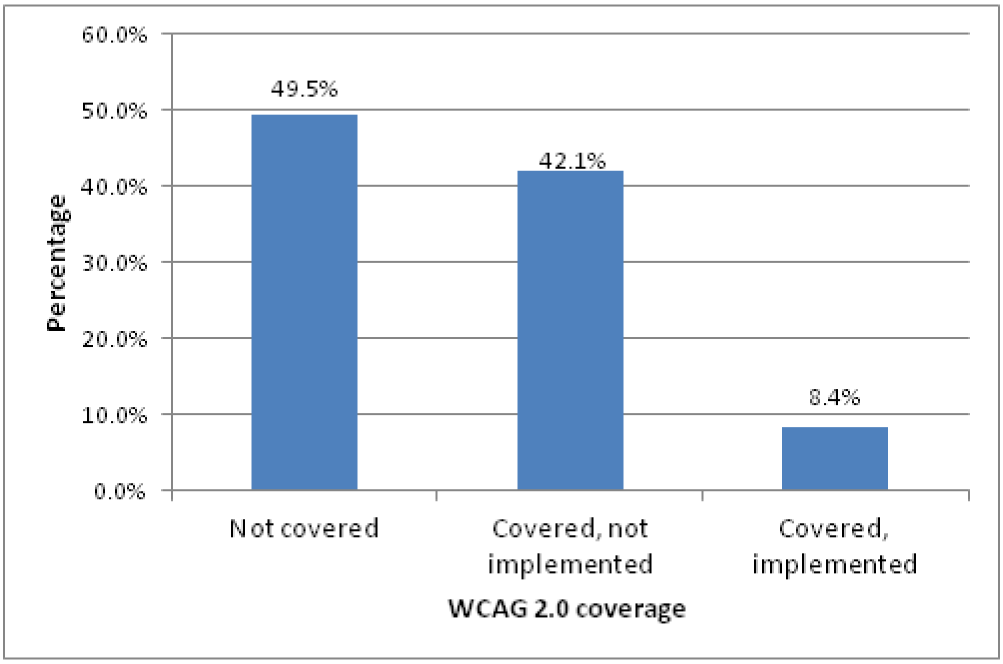


Figure 5.8 –Categories of problems encountered by blind users divided by relevance of WCAG 2.0 SCs and implementation.

For WCAG 2.0, **Error! Reference source not found.** shows a similar breakdown of problems encountered by blind users into categories of relevance and implementation of SCs. The total percentage of problems encountered by blind users that were covered by SCs was 50.5% (the sum of bars 2 and 3 in **Error! Reference source not found.**) and a similarly small percentage of these were implemented by developers (8.4% of all problems encountered by blind users, bar 3 in **Error! Reference source not found.**, or 16.63% of all problems encountered by blind users covered by WCAG 2.0). This means that for WCAG 2.0, the current set of guidelines for web accessibility, almost half of the problems encountered by blind users on websites are not covered.

When comparing the distribution of problems in terms of coverage and implementation, a Related Samples Wilcoxon Signed Rank Test showed there was no significant difference in the coverage of user problems between WCAG 1.0 and WCAG 2.0 across the three relevance and implementation categories ($W_+ = 1.5$, df = 1, p = 1.0).

5.3.2 Problems encountered by partially sighted users and coverage by WCAG checkpoints/success criteria

For WCAG 1.0, **Error! Reference source not found.** shows the breakdown of problems encountered by partially sighted users into categories of relevance of CPs and whether those CPs have been implemented on the website where the problem was encountered. The total percentage of user problems that were covered by CPs was 24.4% (the sum of bars 2 and 3 in **Error! Reference source not found.**) and only a small percentage of those were implemented by developers (7.7% of all problems encountered by partially sighted users, bar 3 in **Error! Reference source not found.**, or 31.56% of all problems encountered by partially sighted users covered by WCAG 1.0). This means that of the problems encountered by partially sighted users on websites, a substantial 75.6% of problems were not covered by WCAG 1.0.

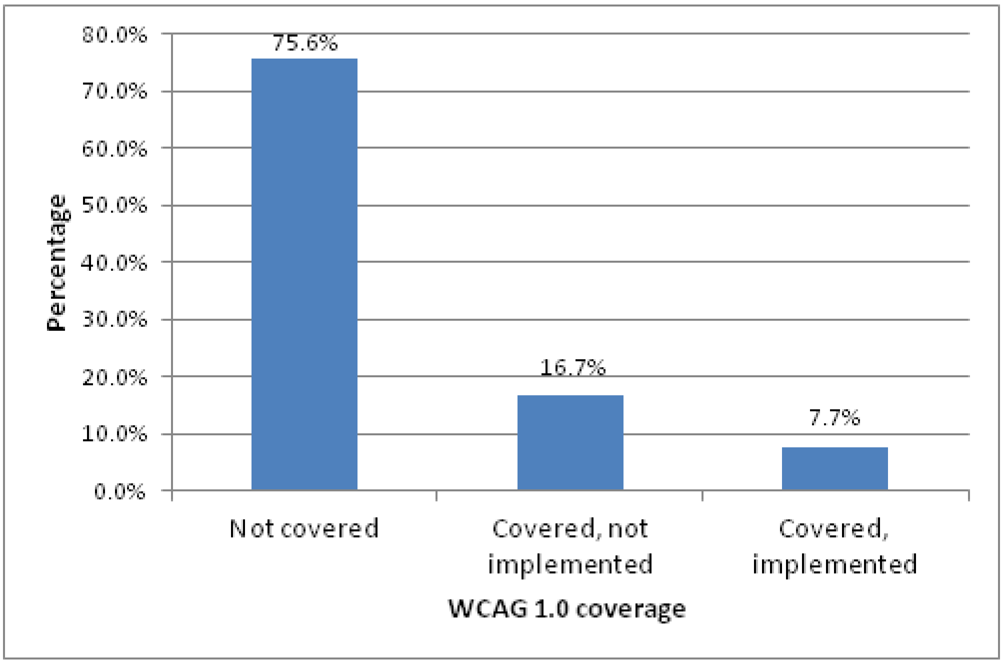


Figure 5.9 –Categories of problems encountered by partially sighted users divided by relevance of WCAG 1.0 CPs and implementation.

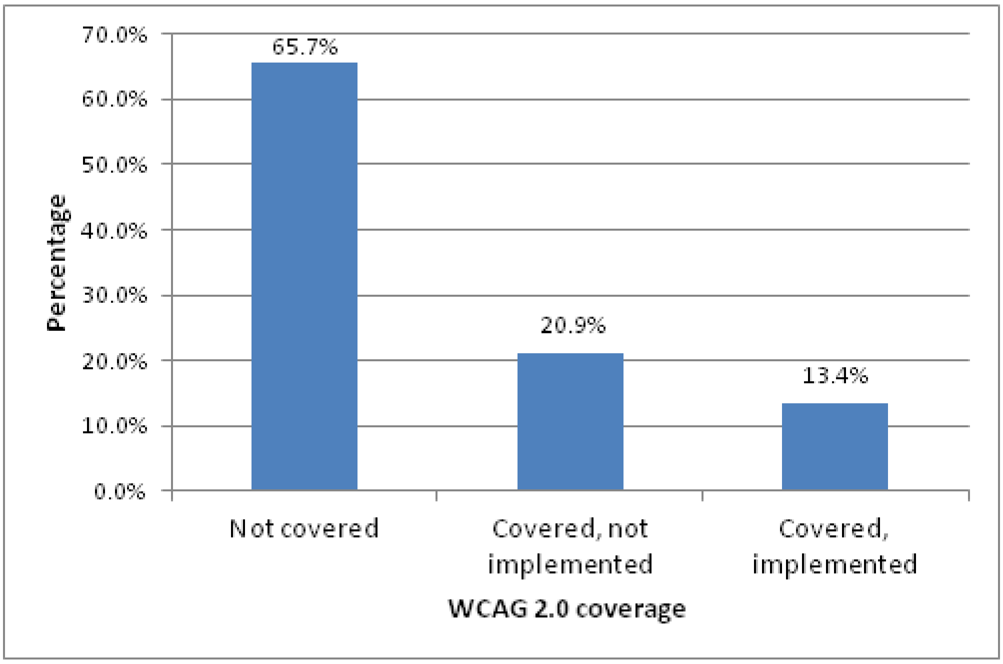


Figure 5.10 –Categories of problems encountered by partially sighted users divided by relevance of WCAG 2.0 SCs and implementation.

For WCAG 2.0, **Error! Reference source not found.** shows a similar breakdown of problems encountered by partially sighted users into categories of relevance and implementation of SCs. The total percentage of problems encountered by partially sighted users that were covered by SCs was 34.3% (the sum of bars 2 and 3 in **Error! Reference source not found.**) and a similarly small percentage of these were implemented by developers (13.4% of all problems encountered by partially sighted users, bar 3 in **Error! Reference source not found.**, or 39.07% of all problems encountered by blind users covered by WCAG 2.0). This means that for WCAG 2.0, a substantial percentage of 65.7% of problems encountered by partially sighted users were not covered.

When comparing the distribution of problems in terms of coverage and implementation, a Related Samples Wilcoxon Signed Rank Test showed there was no significant difference in the coverage of problems encountered by partially sighted users between WCAG 1.0 and WCAG 2.0 across the three relevance and implementation categories ($W_+ = 3$, df = 1, p = 1.0).

5.3.3 Problems encountered by dyslexic users and coverage by WCAG checkpoints/success criteria

For WCAG 1.0, **Error! Reference source not found.** shows the breakdown of problems encountered by dyslexic users into categories of relevance of CPs and whether those CPs have been implemented on the website where the problem was encountered. The total percentage of user problems that were covered by CPs was 7% (the sum of bars 2 and 3 in **Error! Reference source not found.**) and only a small percentage of those were implemented by developers (1.2% of all problems encountered by dyslexic users, bar 3 in **Error! Reference source not found.**, or 17.14% of all problems encountered by dyslexic users covered by WCAG 1.0). This means that of the problems encountered by dyslexic users on websites, a substantial 93.1% of problems were not covered by WCAG 1.0.

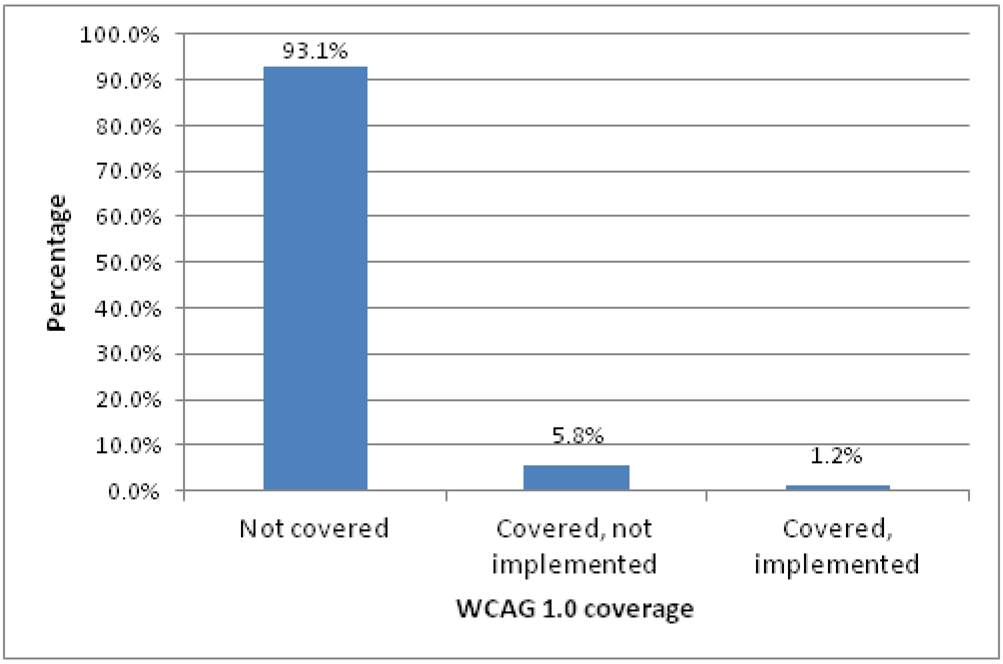


Figure 5.11 –Categories of problems encountered by dyslexic users divided by relevance of WCAG 1.0 CPs and implementation.

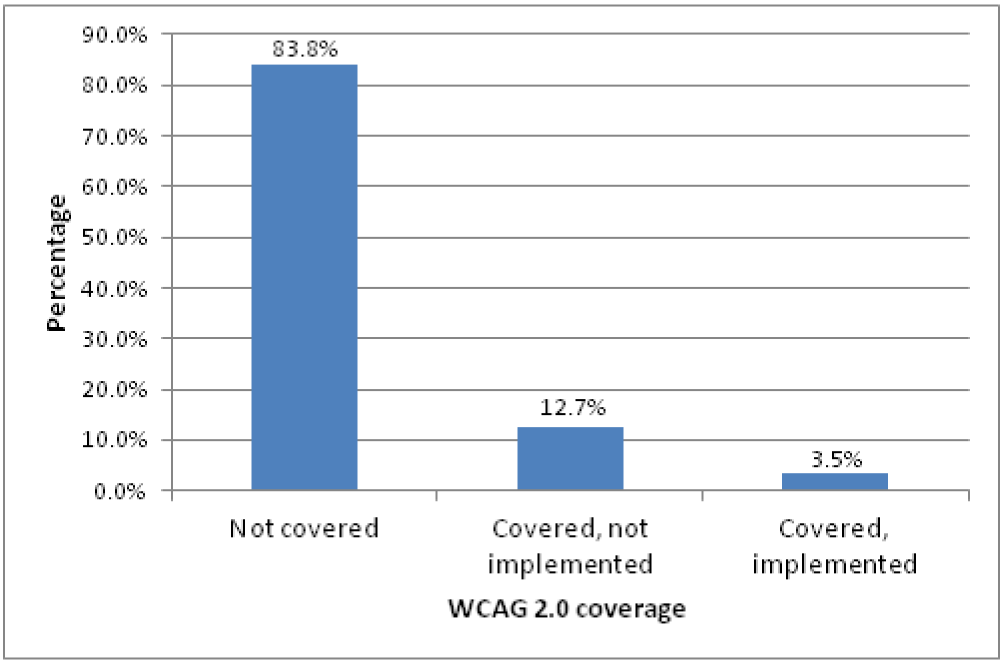


Figure 5.12 –Categories of problems encountered by dyslexic users divided by relevance of WCAG 2.0 SCs and implementation.

For WCAG 2.0, **Error! Reference source not found.** shows a similar breakdown of problems encountered by dyslexic users into categories of relevance and implementation of SCs. The total percentage of problems encountered by dyslexic users that were covered by SCs was 16.2% (the sum of bars 2 and 3 in **Error!**

Reference source not found.) and a similarly small percentage of these were implemented by developers (3.5% of all problems encountered by dyslexic users, bar 3 in **Error! Reference source not found.**, or 21.6% of all problems encountered by blind users covered by WCAG 2.0). This means that for WCAG 2.0, a substantial percentage of 83.8% of problems encountered by dyslexic users were not covered.

When comparing the distribution of problems in terms of coverage and implementation, a Related Samples Wilcoxon Signed Rank Test showed there was no significant difference in the coverage of problems encountered by dyslexic users between WCAG 1.0 and WCAG 2.0 across the three relevance and implementation categories ($W_+ = 0$, df = 1, $p = 1.0$).

5.4 User problems not covered by WCAG

This section presents a list of the main types of problems encountered by blind, partially sighted and dyslexic users that were not covered by WCAG 1.0 or WCAG 2.0. For those problems, one or more relevant CP/SC was either not found, or related CPs/SCs did not cover the nature of the user problems in question in its entirety.

5.4.1 Problems encountered by blind users not covered by WCAG

With respect to types of problems encountered by blind users, 45 of the 65 subcategories of problems (69.2%) had instances of problems not covered by WCAG 1.0, and 36 of 64 (56.3%) had instances of problems not covered by WCAG 2.0.

Regarding the percentage of problems in each subcategories covered by different versions of the guidelines, a Related-Samples Wilcoxon test showed a significant difference between the percentage of problems covered by WCAG 2.0 and WCAG 1.0 in the different subcategories of problems ($W_+ = 2.5$, $N = 64$, $p < 0.001$).

Table 5.2 presents a list of subcategories of problems encountered by partially sighted users with instances not covered by WCAG 1.0 and WCAG 2.0 in ascendant order of coverage by WCAG 2.0. The list contains subcategories with at least 10 problem instances with less than 50% of problems covered by either version of the guidelines.

Table 5.2. Subcategories of problems encountered by blind users with less than 50% of instances covered by WCAG 1.0 or WCAG 2.0 with total number of problems and percentage (number) of problems covered

	Instances of user problems		
Subcategory description	**Total**	**Covered by WCAG 2.0 – % (N)**	**Covered by WCAG 1.0 – % (N)**
Navigation elements do not help users find what they are seeking (Navigation)	99	0	0
Content not found in pages where expected by users (Content)	88	0	0
System too slow (Underlying system characteristics)	27	0	0
No obvious way to return to homepage (Navigation)	21	0	0
No alternative to text in specific format (Text)	17	0	0
Too much information on page (Content)	15	0	0
Information architecture too complex (Information architecture)	15	0	0
Broken link (Underlying system characteristics)	10	0	0
No/insufficient feedback to inform that action has had an effect (Controls, forms or functionality)	72	1.39 (1)	0
Expected functionality not present (Controls, forms or functionality)	31	6.45 (2)	0
Users cannot understand sequence of interaction (Controls, forms or functionality)	14	14.29 (2)	7.14 (1)
Functionality does not work (as expected) (Controls, forms or functionality)	48	18.75 (9)	10.42 (5)
Organisation of content is inconsistent with web conventions/common sense (Content)	39	25.64 (10)	25.64 (10)
Irrelevant content before task content (Content)	87	31.03 (27)	22.99 (20)
Information implied by web page structure not present in page (Content)	12	33.33 (4)	33.33 (4)
Link grouping poor (Links)	11	36.36 (4)	36.36 (4)
Difficult to scan page for specific item (All media types)	18	44.44 (8)	61.11 (11)

As can be seen in Table 5.2, eight subcategories with more than 10 problem categories had all of their problems not covered by WCAG 1.0 and WCAG 2.0. It is noteworthy that two of those subcategories accounted for 13.5% of all problems encountered by blind users.

Neither WCAG 1.0 nor WCAG 2.0 covered problems in the subcategory "*navigation elements do not help users find what they are seeking*". Those problems had CPs/SCs that did not cover the nature of the problems encountered by users in their entirety. WCAG 2.0 SC 2.4.4, for example, states that "*the purpose of each link can be determined from the link text alone or from the link text together with its programmatically determined link context, except where the purpose of the link would be ambiguous to users in general*" (Caldwell et al. 2008). In one example, users were seeking the name of a cabinet minister in charge of public health in the Department of Health. The navigation several options that seemed to be plausible, such as "Public Health", "About us", "Contact", but users could not be sure which one to follow. In such case, the content they were looking for was under "About us". Following SC 2.4.4, fixing the description of each individual link would not necessarily avoid the problem encountered by users, as the main cause of the problem was linked to poor information architecture on the website.

Similarly, problems in the subcategory "*content not found in pages where expected by users*" had CPs/SCs that did not cover the nature of user problems in their entirety, such as SC 2.4.4. Poor information architecture was connected to the main causes of such problems, and not the description of the purpose of individual links.

None of the problem instances of the subcategory "*no obvious way to return to homepage*" were covered by WCAG. Guidelines in both versions have recommendations for consistent navigation, and providing different ways to arrive at different pages, but none state that there should always be ways to return to the home page.

Problems in the subcategory "*no alternative to text in specific format*" were related to the lack of alternative to documents in PDF format. In WCAG 2.0, several techniques were made available to satisfy specific success criteria in PDF documents (Web Accessibility Initiative 2012). This means that WCAG 2.0 recommends that PDF documents be made accessible, and not that alternatives in different formats be provided. In WCAG 1.0, checkpoint 11.1 recommends "*use W3C technologies when they are available and appropriate for a task and use the latest versions when supported*", which is somewhat related to the issue, as the PDF format was not a W3C recommended technology at the point where WCAG 1.0 was published. However, this checkpoint does not provide enough detail of how to solve the issue raised by users.

Issues with "*too much content on pages*" also only had related SCs and CPs that did not cover the problems in their entirety. In WCAG 1.0, CP 12.3 states recommends to "*divide large blocks of information into more manageable groups where natural and appropriate*", and in WCAG 2.0, SC 2.4.10 recommends that *"section headings are* *used to organize the content*". Both recommendations are related to how to organise larger blocks of information, but do not tackle directly the issue of having too much information as a problem itself.

Problems in the subcategory "*information architecture too complex*" represented mainly problems where there were too many steps in navigation to get to a page. This was not covered by either version of WCAG.

Two subcategories related to functional aspects had none of their instances covered by WCAG 1.0 or WCAG 2.0 – "*broken link*" and "*system too slow*" (when pages were too slow to load).

WCAG 2.0 included new success criteria to help identify errors in input in guideline 3.3 – "input assistance: help users avoid and correct mistakes". Only one user problem in the subcategory "*no/insufficient feedback to inform that action has had an effect*", related to the lack of feedback for an empty input field, was covered by SCs related to this guideline. All other problems in this subcategory were related to other types of feedback about application-specific issues that were not listed by guideline 3.3 in WCAG 2.0.

Most problems in the subcategory "*expected functionality not present*" were not covered by WCAG, including issues like the lack of functionality to clean up form fields that pre-defined default values, and not having a search feature. In fact, WCAG 2.0 SC 2.4.5 – "*more than one way is available to locate a Web page*" includes one technique that recommends that a search feature be implemented, but it is one of six options that can be chosen by developers, and hence, not mandatory.

Some problems in the subcategory "*users cannot understand sequence of interaction*" were related to instructions and labels to functionality that was not laid-out in an accessible form to screen readers, and were covered by WCAG. However, in other cases, even though labels and instructions were accessible to users, there was a lack of a proper design of the dialog of the overall interface for a given functionality. This specific issue was not covered by WCAG.

In subcategory "*organisation of content is inconsistent with web conventions/common sense*", some problems were related to elements not being organised logically in groups, such search results not being organised as a properly marked-up list. Such problems were covered by WCAG SCs/CPs. However, many other problems were related to poor disposition of elements on web pages, especially with related information not being placed near each other. WCAG 1.0 and WCAG 2.0 did not cover those cases.

Neither WCAG 1.0 nor WCAG 2.0 covered the majority of problems in the subcategory "*irrelevant content before task content*", with 87 instances. WCAG 2.0 SC 2.4.1 recommends the use of "mechanisms to bypass blocks of content that are repeated on multiple web pages", and WCAG 1.0 CP 13.6 recommended the implementation of bypassing mechanisms for groups of related links. In most problems encountered by users, the problems with irrelevant content before task content were not related to repeated content or to groups of links.

In the subcategory "*information implied by web page structure not present in page*", users encountered headings with no content under it. The content would only appear by clicking on the headings. There were SCs/CPs that were related, but guidelines related to headings did not cover the issue of having headings with no content under them.

Regarding the subcategory "*link grouping poor*", covered problems were related to links that were in separate columns or not organised as a navigation menu, as recommended by the guidelines. Other problems that were not covered were related with badly ordered items in a menu. In one example, a link to bus fares was the last link in a set after a long list of names of bus routes.

In the subcategory "*difficult to scan page for specific item*", covered problems were related to lack of headings to skim past content or to important content that was only in the end of paragraphs. The latter issue was only covered by WCAG 1.0, and not by WCAG 2.0. Other problems were not covered by either version, such as issues with links in the middle of paragraphs, which users found difficult to scan when reading text.

5.4.2 Problems encountered by partially sighted users not covered by WCAG

With respect to types of problems encountered by partially sighted users, 29 of the 55 subcategories of problems (52.72%) had instances of problems not covered by WCAG 1.0, and 21 of 55 (38.18%) had instances of problems not covered by WCAG 2.0.

Regarding the percentage of problems in each subcategory covered by different versions of the guidelines, a Related-Samples Wilcoxon test showed a significant difference between the percentage of problems covered by WCAG 2.0 and WCAG 1.0 in the different subcategories of problems ($W+ = 3.069$, $N = 55$, $p < 0.002$).

Table 5.3 presents a list of subcategories of problems encountered by partially sighted users with instances not covered by WCAG 1.0 and WCAG 2.0 in ascendant order of coverage by WCAG 2.0. The list contains subcategories with at least 10 problem instances with less than 50% of problems covered by either version of the guidelines.

As can be seen in Table 5.3, seven subcategories had none of their problem instances covered by either WCAG 1.0 or WCAG 2.0. Similarly to the situation explained for blind users, subcategories "*navigation elements do not help users find what they are seeking*" and "*content not found in pages where expected by users*" had SCs/CPs that did not cover to the nature of the problems encountered by users in their entirety. The same also applied to subcategories "*too much information on page*", "*no obvious way to return to homepage*", "*irrelevant content before task content*", "*information architecture too complex*" and "*expected functionality not present*", for the same reasons that problems in this subcategory encountered by blind users were not covered by WCAG.

In the subcategory "*difficult to scan page for specific item*", three problem instances were covered by WCAG 1.0 CP 3.18 - "*place distinguishing information at the beginning of headings, paragraphs, lists, etc*", as the problems were related to important information placed at the end of paragraphs. However, the remainder 93.18% of problems in this subcategory had checkpoints that had some relation to the problem, but did not cover them entirely. This was no longer a requirement in WCAG 2.0.

Regarding problems in the subcategory "No/insufficient feedback to inform that action has had an effect", only one problem was related to content added dynamically on a page via script not perceived by the user, which was covered by WCAG. Other issues were related to poor feedback messages or lack of any feedback.

Table 5.3. Subcategories of problems encountered by partially sighted users with less than 50% of instances covered by WCAG 1.0 or WCAG 2.0 with total number of problems and percentage (number) of problems covered

	Instances of user problems		
Subcategory description	**Total**	**Covered by WCAG 2.0 – % (N)**	**Covered by WCAG 1.0 – % (N)**
Navigation elements do not help users find what they are seeking (Navigation)	78	0	0
Content not found in pages where expected by users (Content)	77	0	0
Irrelevant content before task content (Content)	40	0	0
Too much information on page (Content)	33	0	0
Expected functionality not present (Controls, forms or functionality)	33	0	0
No obvious way to return to homepage (Navigation)	13	0	0
Information architecture too complex (Information architecture)	12	0	0
Difficult to scan page for specific item (All media types)	44	0	6.82 (3)
No/insufficient feedback to inform that action has had an effect (Controls, forms or functionality)	15	6.67 (1)	6.67 (1)
Functionality does not work (as expected) (Controls, forms or functionality)	30	6.67 (2)	10 (3)
Inability to change presentation of audio, video or multimedia (Audio, video or multimedia)	31	16.13 (5)	6.45 (2)
Organisation of content is inconsistent with web conventions/common sense (Content)	29	17.24 (5)	20.69 (6)
Default presentation of control or form element not adequate (Controls, forms or functionality)	43	30.23 (13)	53.49 (23)
Default presentation of image not adequate (Images)	32	37.5 (12)	31.25 (10)
Users cannot understand sequence of interaction (Controls, forms or functionality)	10	40 (4)	0
Users cannot make sense of content (Content)	28	42.86 (12)	14.29 (4)

In the subcategory "functionality does not work", some problems were related to non-standard HTML components (such as select boxes) implemented with scripts that did not work properly with screen magnifiers, which were covered. Other issues were related to malfunctioning of implemented features, especially ineffective searches, and those were not covered by the guidelines.

Some problems in the subcategory “*inability to change presentation of audio, video or multimedia*” were related to the inability to pause animations or moving content, and were covered by WCAG 2.0 SC 2.2.2 – “*[allow] pause, stop, hide: For moving, blinking, scrolling, or auto-updating information*...” or WCAG 1.0 CP 7.3 – “..*avoid movement in pages*”. However, other problems in this category were not covered. Those problems were mainly related to not being able to change the speed of the presentation of multimedia content, such as video. In those cases, users did not want just to pause every time they wanted to read something, but to make the video play more slowly.

In the subcategory “*organisation of content is inconsistent with web conventions/common sense*”, some problems were related to navigation bars placed inconsistently in different pages, which was covered by WCAG 1.0 and 2.0. Most problems in this subcategory, however, were related to poor visual disposition of content, such as related content not placed in adjacent places. Such issues were not covered by the guidelines.

It is noteworthy that only a small proportion of problems encountered by partially sighted users in subcategories “*default presentation of control or form element not adequate*” and “*default presentation of control or form element not adequate*” were covered by WCAG 1.0 or WCAG 2.0. Many such problems were related to size or colour contrast of images or controls that were not text. WCAG 1.0 CP 2.2 recommends to “*ensure that foreground and background colour combinations provide sufficient contrast when viewed by someone having colour deficits or when viewed on a black and white screen*”, applicable to both text and images. In WCAG 2.0, on the other hand, SCs 1.4.3 and 1.4.6 only apply minimum contrast ratios to “*the visual presentation of text and images of text*”. This means that the referred WCAG 2.0 SC would not be applicable to the colour contrast of a non-textual graphical icon in a button, for example.

Like for blind and dyslexic users, most problems in the subcategory “*users cannot understand sequence of interaction*“ were caused by the lack of a proper design of the dialog of the overall interface for a given functionality and not by problems with labels of individual controls.

In subcategory “*users cannot make sense of content*”, many problems were not covered by either WCAG 1.0 or WCAG 2.0. Such problems were not related to the use of jargon, acronyms or the level of the language, but to incomplete information in websites. For example, users found it difficult to work out the posting price scheme due to the lack of explanations.

5.4.3 Problems encountered by dyslexic users not covered by WCAG

With respect to types of problems encountered by dyslexic users, 32 of the 44 subcategories of problems (72.7%) had instances of problems not covered by WCAG 1.0, and 27 of 44 (61.4%) had instances of problems not covered by WCAG 2.0.

Regarding the percentage of problems in each subcategories covered by different versions of the guidelines, a Related-Samples Wilcoxon test showed a significant difference between the percentage of problems covered by WCAG 2.0 and WCAG 1.0 in the different subcategories of problems ($W+ = 2.836$, $N = 44$, $p < 0.005$).

Table 5.4 presents a list of subcategories of problems encountered by dyslexic users with instances not covered by WCAG 1.0 and WCAG 2.0 in ascendant order of coverage by WCAG 2.0. The list contains subcategories with at least 10 problem instances with less than 50% of problems covered by either version of the guidelines.

As can be seen in Table 5.4, 10 subcategories of problems encountered by dyslexic users had none of their instances covered by either WCAG 1.0 or WCAG 2.0. Those problems account for more than 50% of all problems encountered by dyslexic users.

As explained in Section 5.4.1 for blind users, problems in subcategories "*content not found in pages where expected by users*" and "*navigation elements do not help users find what they are seeking*" were not covered by WCAG SCs/CPs. Guidelines related to the description of the purpose of an individual link had only a marginal relation to the problems, as they did not address the underlying issue of poor information architecture in the websites.

Table 5.4. Subcategories of problems encountered by dyslexic users with less than 50% of instances covered by WCAG 1.0 or WCAG 2.0 with total number of problems and percentage (number) of problems covered

Subcategory description	Instances of user problems		
	Total	**Covered by WCAG 2.0 – % (N)**	**Covered by WCAG 1.0 – % (N)**
Content not found in pages where expected by users (Content)	112	0	0
Navigation elements do not help users find what they are seeking (Navigation)	87	0	0
Too much information on page (Content)	34	0	0
Organisation of content is inconsistent with web conventions/common sense (Content)	30	0	0
Functionality does not work (as expected) (Controls, forms or functionality)	29	0	0
Users cannot understand sequence of interaction (Controls, forms or functionality)	17	0	0
No/insufficient feedback to inform that action has had an effect (Controls, forms or functionality)	17	0	0
Irrelevant content before task content (Content)	13	0	0
Users inferred the existence of functionality where there was not one (Controls, forms or functionality)	13	0	0
Navigation bar not salient (Navigation)	10	0	0
Expected functionality not present (Controls, forms or functionality)	37	0	2.7 (1)
Difficult to scan page for specific item (All media types)	72	5.56 (4)	5.56 (4)
Destination not what was anticipated (Navigation)	11	36.36 (4)	36.36 (4)
Users cannot make sense of content (Content)	29	37.93 (11)	6.9 (2)
Default presentation of text not adequate (Text)	44	43.18 (19)	0

Some issues related to "*too much content on page*" and "*organisation of content is inconsistent with web conventions/common sense*" had SCs/CPs that were related to the issues but did not cover them entirely. However, the problems with limiting the amount of content in a page and the visual disposition of content that caused the problems reported by users were not directly addressed by any guideline in WCAG.

Problems with malfunctioning of functionality in subcategory "*functionality does not work (as expected)*", especially with search features that do not find expected content, were not covered by either WCAG 1.0 or WCAG 2.0.

Problems in the subcategory "*users cannot understand sequence of interaction*" had a limited relation to WCAG 2.0 SC 3.3.2 – "*labels or instructions are provided when content requires user input*". However, many problems were not caused by the lack of labels or instructions to individual interface components, but by lack of a proper design of the dialog of the overall interface for a given functionality.

In the subcategory "*no/insufficient feedback to inform that action has had an effect*", problems with lack of feedback were not covered by either WCAG 1.0 or WCAG 2.0. None of the problems with feedback were classified as "input error [that] is automatically detected", as described in WCAG 2.0 SCs 3.3.1 and 3.3.3.

Similarly to the problems encountered by blind and partially sighted users in the subcategory "*irrelevant content before task content*", problems encountered by dyslexic users were not covered by WCAG, as the content they considered irrelevant was not repeated content from other pages, as stated by WCAG 2.0 SC 2.4.1 and WCAG 1.0 CP 13.6.

Regarding problems in "*users inferred the existence of functionality where there was not one*", where users thought that a text in bold that looked like a link but was not one, for example, no mention of this kind of problem was found in WCAG 1.0 and WCAG 2.0.

Problems with "*navigation bar not salient*" were not fully covered by WCAG 1.0 or WCAG 2.0. Some SCs/CPs had some relation to those issues, such as WCAG 1.0 CP 13..5 – "*provide navigation bars to highlight and give access to the navigation mechanism*" and WCAG 2.0 SC 2.4.5 – "*more than one way is available to locate a Web page within a set of Web pages*", which includes providing navigation bars as a possibility. However, neither of them provides guidance to make navigation bars salient on the screen.

In the subcategory "*expected functionality not present*", only one of the 37 problem instances was covered by WCAG 1.0 CP 13.7 – "*if search functions are provided, enable different types of searches for different skill levels and preferences*". Neither WCAG 1.0 nor WCAG 2.0 covered other problems with features expected by users not present, such as auto-complete for users with spelling difficulties.

The majority of problems in the subcategory "*difficult to scan page for specific item*" were not covered by WCAG. In only four of 72 instances were problems related specifically to the lack of headings to break large amounts of content into smaller groups, as stated by WCAG 1.0 CP 12.3 – "*divide large blocks of information into more* *manageable groups where natural and appropriate*" and WCAG 2.0 SC "*section headings are used to organize the content*".

While some problems in the subcategory "*destination not what was anticipated*" were caused by issues covered by WCAG, such as unclear link destination, other problems were not covered by the guidelines. In many cases, the content of a certain section was in a separate website (such as NHS Direct and NHS Choices). Many problems were caused by users expecting to return to the website where they started from, but arriving at the home page of a distinct website. This issue was not fully covered by guidelines related to link destination, as the purpose of the link "Home" was clear within the context of a web page, but not in the context of users having been redirected from a different website. WCAG 2.0 SC 3.2.5 has some relation with the problem, as it states that "changes of context are initiated only by user request". However, none of the techniques currently available address the issue of changing to different related websites.

In the subcategory "*users cannot make sense of content*", some problems were covered by guidelines that dealt with jargon, difficult words and level of language. Other issues with incomplete information about the specific domain of the website were not covered.

The subcategory "*default presentation of text not adequate*" had significantly more problems covered by WCAG 2.0 than WCAG 1.0. WCAG 2.0 SC 1.4.8 covered issues related to providing mechanisms to change the presentation of blocks of text, including changes in foreground and background colour, character width, paragraph justification, line spacing and font size. However, many other issues with presentation of text encountered by dyslexic users were still not covered. The guidelines state that users should be able to change font size. Many users reported problems that they considered the default size too small in spite of being able to change the size. Other issues not covered by the guidelines include problems with text in italics, inappropriate typeface and text in columns.

5.5 User problems related to WCAG guidelines successfully implemented on websites

In order to understand why WCAG 2.0 SCs or WCAG 1.0 CPs do not solve some problems encountered by users, analyses were performed on problems encountered on web pages where directly relevant SCs/CPs were implemented and yet users still had problems.

In the following sections, the main types of problems where implemented CPs/SCs did not solve users' problems are presented for blind, partially sighted and dyslexic users, respectively.

5.5.1 Problems encountered by blind users related to WCAG guidelines successfully implemented on websites

Table 5.5 presents the subcategories of problems encountered by blind users where at least 3 instances of problems were related to interface components that met the criteria to pass one or more directly relevant CPs/SCs, and users still encountered problems. The table is ordered in descending order of the percentage of problems of each subcategory that were covered and implemented by WCAG 2.0.

For blind users, 5.6% of all user problems or 13.02% of all user problems covered by WCAG 1.0 were encountered in web pages that met the criteria to pass WCAG CPs, and for WCAG 2.0, this corresponded to 8.4% user problems, or 16.63% of all user problems covered by WCAG 2.0.

As can be seen in Table 5.5, the subcategory "heading structure violated" had 100% of its problem instances covered and implemented by WCAG 2.0 SCs. Users reported problems when the heading structure was not logical to them, such as having a level-2 heading as the first heading on a page, without a preceding level-1 heading, or jumping from a level-1 heading straight to a level-3 heading, without a heading at level 2 in the middle. Implementing headings correctly is covered by WCAG 2.0 SC 1.3.1 – "*information, structure, and relationships conveyed through presentation can be programmatically determined or are available in text*", and more specifically by technique H42 – "*using h1-h6 to identify headings*". In fact, in one of the examples of correct implementations of H42, there is an excerpt of HTML code where the heading order is H2 – H1. For WCAG 1.0, on the other hand, violating the heading structure is considered a violation of CP 3.5 – "*use header elements to convey document structure and use them according to specification*". In fact, WCAG 1.0 techniques for headings (Chisholm et al. 2000) explicitly state that "... *in HTML, H2 elements should follow H1 elements, H3 elements should follow H2 elements, etc. Content developers should not "skip" levels (e.g., H1 directly to H3)*".

Table 5.5. Subcategories of problems encountered by blind users with their total number of problems and percentage (number) of problems covered by WCAG 2.0 SCs or WCAG 1.0 CPs and implemented

Subcategory description	Total user problems	Covered and implemented % (N)	
		WCAG 2.0	WCAG 1.0
Heading structure violated (Headings)	9	100 (9)	0
No alternative to information presented in tables (Tables)	12	100 (12)	100 (12)
Meaning in content is lost or modified due to transformation (Content)	8	62.5 (5)	0
No enhancement to audio, video or multimedia (Audio, video or multimedia)	31	51.61 (16)	0
Link destination not clear (Links)	117	38.46 (45)	10.26 (12)
Language too complicated for perceived target audience (Content)	9	33.33 (3)	33.33 (3)
Users cannot make sense of content (Content)	66	6.06 (4)	3.03 (2)
It is not clear what particular controls or form elements do (Controls, forms or functionality)	79	5.06 (4)	0
Irrelevant content before task content (Content)	87	4.6 (4)	4.6 (4)

The subcategory "*no alternative to information presented in tables*" had 100% of its instances covered and implemented for both WCAG 1.0 and 2.0. WCAG 2.0 SC 1.3.1 provides several mandatory techniques to make tables accessible, such as providing captions (technique H39), using the summary attribute to explain how a table works (technique H73) and using the scope attribute (technique H63) and cell/header ids (H43) to relate table cells to their headings. Similarly, WCAG 1.0 has six checkpoints related to guideline 5 – "*create tables that transform gracefully*" that cover recommendations to make tables more accessible. However, in the problems reported by users when there was no alternative available, the referred tables had successfully implemented all the techniques recommended by the guidelines. Nevertheless, it was observed that manipulating tables per se was considered a very complex task for users who had limited experience with specific screen reader features to read tables.

In the subcategory "*meaning in content is lost or modified due to transformation*", five instances were covered and implemented WCAG 2.0 SC 3.1.5 – "*when text requires reading ability more advanced than the lower secondary education level [...] a version that does not require reading ability more advanced than the lower secondary education level, is available*". In those problems users found it difficult to read a legal text with difficult language, and resorted to a simplified summary to try and find a date in the document without success, as the date had not been included in the summary. Technique G86, which is sufficient for SC 3.1.5, provides guidance to providing simplified summaries. However, the test procedure of such technique only requires to "*1. measure the readability of the summary, and 2) check that the summary requires reading ability less advanced than the lower secondary education level.*".

In the subcategory "*language too complicated for perceived target audience*", users who encountered difficult language in a legal website did not even find a simplified summary at the end of the page, despite it being sign-posted according to WCAG recommendations.

In most problems in the subcategory "*no enhancement to audio, video or multimedia*", blind users reported problems with the lack of audio descriptions of video content. For a page containing pre-recorded videos, it passes WCAG 2.0 SC 1.2.3 (Level A), if it provides an audio description or another alternative for all these videos. The only other alternative mentioned by WCAG is a text description of the videos with a text transcript of the audio tracks, all indexed by time. However, in a somewhat complex relationship between SC 1.2.3 and SC 1.2.5, if audio description (as opposed to the text description) is provided for all pre-recorded videos, the page also passes SC 1.2.5 (Level AA). As shown in Table 5.5, there were 31 problems in this subcategory. Of those problems, 51.61% of the web pages passed SC 1.2.3 at Level A by providing an appropriate text description. These problems were covered by WCAG 2.0 and were implemented correctly, but users rejected that implementation because they wanted an audio description. The remaining problems in the category were covered by WCAG 2.0, but not implemented properly because there was no audio description or any other alternative provided.

The largest number of user problems in Table 5.5 are in the subcategory "*link destination not clear"*, which accounted for 8.46% of all problems encountered by blind users. In 38.46% of problems in this category, the website had properly implemented

WCAG 2.0 SC 2.4.4 (level A) regarding the description of link purpose, and yet users still had problems determining where the links lead. Unlike SC 2.4.9 (level AAA), which states that the link purpose should be defined by the link in its own, SC 2.4.4 allows for the link purpose to be determined "in context", which includes a preceding heading, surrounding paragraph, list element or table, for example. Although passing SC 2.4.4, many users still encountered problems when reading links out of context in a "links list" feature, commonly used by screen-reader users. In WCAG 1.0, on the other hand, CP 13.1 makes an exception for links that are in a sequence, such as page numbers in search results. Although passing WCAG 1.0 CP 13.1, those links were also reported as problematic by users.

In the subcategory "*users cannot make sense of content*", users encountered abbreviations that they did not understand followed by the description. Although having a description following the acronym was sufficient to pass WCAG 2.0, for example, many users did not relate the description to the acronym.

In some problems in the subcategory "*it is not clear what particular controls or form elements do*", an input field did not have an associated LABEL element to identify the field. However, one sufficient technique allows "*using an adjacent button to label the purpose of a field*" (G167) to pass WCAG 2.0 SC 3.3.2 "*labels or instructions are provided when content requires user input*". For example, a button named "search" after an input field would identify what the purpose of a field is. However, in three instances, the use of this technique did not avoid problems with users not knowing what the purpose of the field was.

In the subcategory "*irrelevant content before task content*", some websites successfully implemented a link to "skip to content" to enable users to skip past a repeated navigation bar, for example. Even though this feature was implemented, some users did not use it, and reported problems with having too much irrelevant content.

5.5.2 Problems encountered by partially sighted users related to WCAG guidelines successfully implemented on websites

Table 5.6 presents the subcategories of problems encountered by partially sighted users where at least 3 instances of problems were related to interface components that met the criteria to pass one or more directly relevant CPs/SCs, and users still encountered problems. The table is ordered in descending order of the percentage of problems of each subcategory that were covered and implemented by WCAG 2.0.

For partially sighted users, 7.7% of all user problems or 31.56% of all user problems covered by WCAG 1.0 were encountered in web pages that met the criteria to pass WCAG CPs, and for WCAG 2.0, this corresponded to 13.4% of all user problems, or 39.07% of all user problems covered by WCAG 2.0.

Table 5.6. Subcategories of problems encountered by partially sighted users with their total number of problems and percentage (number) of problems covered by WCAG 2.0 SCs or WCAG 1.0 CPs and implemented

Subcategory description	Total user problems	Covered and implemented % (N)	
		WCAG 2.0	WCAG 1.0
No alternative to information presented in tables (Tables)	5	80 (4)	80 (4)
Meaning in content is lost or modified due to transformation (Content)	4	75 (3)	0
Default presentation of text not adequate (Text)	157	55.41 (87)	27.39 (43)
No enhancement to audio, video or multimedia (Audio, video or multimedia)	6	50 (3)	0
Link destination not clear (Links)	13	46.15 (6)	30.77 (4)
Default presentation of image not adequate (Images)	32	21.88 (7)	9.38 (3)
Inability to change presentation of text (Text)	22	13.64 (3)	13.64 (3)
Inability to change presentation of audio, video or multimedia (Audio, video or multimedia)	31	9.68 (3)	0
Default presentation of control or form element not adequate (Controls, forms or functionality)	43	6.98 (3)	23.26 (10)

As seen in Table 5.6, in 80% of problems related to "*no alternative to information presented in tables*", partially sighted users also reported problems with lack of alternatives with tables that had implemented all requirements for accessible tables available in WCAG 1.0 and WCAG 2.0. For partially sighted users who use speech synthesis to read content on web pages, the techniques available in WCAG (such as relating table cells to their corresponding headings) may be of use in the same fashion as for blind users. However, for other users that do not use speech, complex tables may be difficult for those users even if they use correct mark-up. Relating the content of a cell to its headings to make sense of it may involve complex visual operations with content that may not fit all at the same time in users' viewport.

Like for blind users, partially sighted users encountered problems in the subcategory "*meaning in content is lost or modified due to transformation*" that were covered in web pages that implemented WCAG 2.0 SC 3.1.5 – "*when text requires reading ability more advanced than the lower secondary education level [...] a version that does not require reading ability more advanced than the lower secondary education level, is available*". In those problems, although a simplified summary had been provided, key information to users' task had not been provided in the summary. The web pages had passed the tests for Technique G86, as the test procedure only required to *"check that the summary requires reading ability less advanced than the lower secondary education level*".

As shown in Table 5.6, three subcategories related to the inadequacy of the default presentation had problem instances related to web pages that implemented related WCAG SCs/CPs. In the subcategory "*default presentation of text not adequate*", 55.41% of problems implemented relevant WCAG 2.0 SCs and 27.39% implemented relevant WCAG 1.0 CPs. In the subcategory "*default presentation of image not adequate*", the percentages were 21.88% for WCAG 2.0 and 9.38% for WCAG 1.0, and in "*default presentation of control or form element not adequate*", 6.98% for WCAG 2.0 and 23.26% for WCAG 1.0. Some issues were related to the size of the content being considered too small by users, and the large majority was related to poor colour contrast.

The issue with text size is covered by WCAG 2.0 SC 1.4.4 – "*except for captions and images of text, text can be resized without assistive technology up to 200 percent without loss of content or functionality*". In WCAG 1.0, CP 3.4 states to "*use relative rather than absolute units in mark-up language attribute values and style sheet property values*". Both aim at allowing users to resize content according to their needs. In many cases, users were indeed able to resize text in web pages, as techniques for WCAG 2.0 SC 1.4.4 and WCAG 1.0 CP 3.4 were implemented. However, in some situations, the original font size was so small that the required level of magnification (at times much higher than 200%) to enable users to see would make it very difficult to read the page in a small view port, often making users lose the context of other content in the layout.

Regarding issues with colour contrast, they are dealt with differently in WCAG 1.0 and WCAG 2.0. In WCAG 1.0, CP 2.2 states "*ensure that foreground and background color combinations provide sufficient contrast when viewed by someone having colour deficits or when viewed on a black and white screen. [Priority 2 for images, Priority 3 for text].*". Although no specific colour contrast ration was defined in the official recommendations, the most common test for colour contrast used to test WCAG 1.0 was the algorithm defined in the document "*Techniques For Accessibility Evaluation And Repair Tools*" (AERT) (Ridpath and Chisholm 2000). The algorithm calculates two values for "colour brightness" and "colour difference" and proposes minimum values for each of them.

In WCAG 2.0, colour contrast is dealt with in two different SCs: SC 1.4.3 (level AA) stipulates a minimum colour contrast, stating that that "*the visual presentation of text and images of text has a contrast ratio of at least 4.5:1*", with an exception for large text requiring contrast of at least 3:1; SC 1.4.6 (level AA) provides requirements for "enhanced contrast", with contrast ratio of at least 7:1, and 4.5:1 for large text. The definition of the minimum contrast ratio of 3:1 for large text was based on the ISO-9241-3 standard (International Standards Organization 1993). The argument (Cooper et al. 2010a, Vanderheiden 2009) for using the contrast ratio of 4.5:1 in WCAG 2.0 was based on findings that visual acuity of 20/40 is associated with a contrast sensitivity loss of roughly 1.5 (Arditi and Faye 2004), and hence, with the proposed contrast ratio of 4:5 resulting from the minimum ratio of 3:1 times 1.5. The same principle was followed to define that a user with 20/80 visual acuity would require contrast of about 7:1 (Cooper et al. 2010a, Vanderheiden 2009).

It was found in the present study that 70 problem instances covered by WCAG 2.0 and WCAG 1.0 were related to colour contrast. Of those problems, 37.14% were related to web page components that passed the test defined by the AERT algorithm used in many tools for WCAG 1.0 (Ridpath and Chisholm 2000). With WCAG 2.0, a total of 74.29% of the problems were related to web page components that passed WCAG SC related to colour contrast at some level, with 30% passing SC 1.4.3 (level AA) only, and another 44.29% passing both SC 1.4.3 (level AA) and 1.4.6 (level AAA).

In the subcategory "*no enhancement to audio, video or multimedia*", 50% of problems related to the lack of audio description passed WCAG 2.0 SCs. Like blind users, partially sighted users also rejected the implementation of text description of the videos with a text transcript of the audio tracks indexed by time, which was allowed to pass WCAG 2.0 SC 1.2.3.

Like blind users, some partially sighted users also encountered problems in the subcategory "*link destination not clear*" when trying to determine the destination of a link. In such cases, the links were related to exceptions allowed by WCAG 2.0 and WCAG 1.0 in the subcategory. Some links were page numbers in a list of links to pages, and others were links that had to be placed in the context of a heading, for example (a link named "read more" about a news article needs to be related to the preceding heading with the article's name). However, even though those examples would pass WCAG 1.0 CP 13.1 and/or WCAG 2.0 SC 2.4.4, users still found problems. Depending on the magnification level, determining the context in which a link is placed can be very difficult to users due to the very small viewport that they have.

Some problems in the subcategory "*inability to change presentation of text*" were related to users not being able to change the presentation of distorted text presented in "captchas" used as security checks. WCAG 2.0 SC 1.1.1 suggests that whenever captchas are used, an alternative should be provided. In 13.64% of problems in this subcategory, captchas had an alternative with a sound for users to decipher, which would be sufficient to pass WCAG 2.0 SC 1.1.1. However, users found it difficult to use the audio version of the captcha, as there was too much noise, and found that it still did not help them.

In the subcategory "*inability to change presentation of audio, video or multimedia*", users encountered problems where they could not stop an animation or movement on a web page. However, in 9.68% of those cases, the web pages had implemented ways to pause or stop the animation (with a pause button, for example), as suggested by WCAG 2.0 SC 2.2.2: "*for any moving, blinking or scrolling information [...] there is a mechanism for the user to pause, stop, or hide it*". However, due to their limited viewport, users found it difficult to find such buttons and were not able to stop the animation, even though a button was available.

5.5.3 Problems encountered by dyslexic users related to WCAG guidelines successfully implemented on websites

Table 5.7 presents the subcategories of problems encountered by dyslexic users where at least 3 instances of problems were related to interface components that met the criteria to pass one or more directly relevant CPs/SCs, and users still encountered problems. The table is ordered in descending order of the percentage of problems of each subcategory that were covered and implemented by WCAG 2.0.

For dyslexic users, 1.2% of all user problems or 17.14% of all user problems covered by WCAG 1.0 were encountered in web pages that met the criteria to pass WCAG CPs, and for WCAG 2.0, this corresponded to 3.5% of all user problems, or 21.6% of all user problems covered by WCAG 2.0.

Table 5.7. Subcategories of problems encountered by dyslexic users with their total number of problems and number (percentage) of problems covered by WCAG 2.0 SCs or WCAG 1.0 CPs and implemented

Subcategory description	Total user problems	Covered and implemented % (N)	
		WCAG 2.0	WCAG 1.0
Meaning in content is lost or modified due to transformation (Content)	8	87.5 (7)	0
Language too complicated for perceived target audience (Content)	15	26.67 (4)	6.67 (1)
Default presentation of text not adequate (Text)	44	18.18 (8)	0

As seen in Table 5.7, like blind and partially sighted users, dyslexic users also encountered problems in the subcategory "*meaning in content is lost or modified due to transformation*" that were covered in web pages that implemented WCAG 2.0 SC 3.1.5 – "*when text requires reading ability more advanced than the lower secondary education level [...] a version that does not require reading ability more advanced than the lower secondary education level, is available*". In those problems, although a simplified summary had been provided, key information to users' task had not been provided in the summary. The web pages had passed the tests for Technique G86, as the test procedure only required to *"check that the summary requires reading ability less advanced than the lower secondary education level*".

Also similarly to blind users, in the subcategory "*language too complicated for perceived target audience*", users who encountered with difficult language in a legal website did not even find a simplified summary at the end of the page, despite it being sign-posted according to WCAG recommendations.

In the subcategory "Default presentation of text not adequate", 19 of the 44 problem instances were covered by WCAG 2.0. All of those problems were related to SC 1.4.8 (level AAA), related to the visual presentation of blocks of text. This SC has five requirements (Caldwell et al. 2008), the first being to "*ensure foreground and background colours can be selected by the user*". Of the 19 problems covered by WCAG 2.0 SC 1.4.8, 15 were related to colour background or foreground making it difficult to read. In 8 occasions, the web pages where such problems occurred had met the first requirement of SC 1.4.8 by implementing one of the five sufficient techniques available. In all such occasions, websites implemented technique G156 – "*using a technology that has commonly-available user agents that can change the foreground and background of blocks of text*". This technique assumes that users can change background and foreground colours using customisation options available in web browsers. It is worth noting that none of the dyslexic users in the present study used the browser's settings to change foreground and background colour, despite them being advised that they were allowed to make any changes to the settings in the computer that they wanted.

Another technique suggested by WCAG 2.0 to satisfy the first requirement for SC 1.4.8 is G175 - "*providing a multi colour selection tool on the page for foreground and background colours*", in which websites would implement an embedded tool to enable colour selection. However, in the 8 occasions where websites passed this requirement, this technique was not used.

5.6 Severity of user problems and priority of related WCAG checkpoints/success criteria

This section presents the results of the analysis of the correlation between the severity ratings of user problems and the priority levels of WCAG CPs/SCs.

Given that different users could have encountered problems that were caused by the same website problem in different instances, this analysis was based on the mean severity ratings of website problems.

For some particular types of problems, it could be that more than one WCAG CP/SC would be related to the problem encountered by users, and that they could have different priority levels. In such cases, the analysis considered two different approaches: one considering the highest priority level and another considering the mean of the priority levels. In order to calculate this for WCAG 2.0, SCs at level A were assigned priority 1, SCs at level AA were assigned priority 2 and SCs at level AAA were assigned priority 3. Only problems that were covered by WCAG 1.0 or WCAG 2.0 were considered.

The following sections present the analysis of the correlations between the severity ratings of user problems and WCAG 1.0 and 2.0 priority levels for problems encountered by blind, partially sighted and dyslexic users, respectively.

5.6.1 Severity of problems encountered by blind users and priority of related WCAG checkpoints/success criteria

For blind users, 400 distinct website problems were covered by WCAG 2.0. No significant correlation was found between the highest priority levels of related SCs and the mean severity ratings of user problems ($r = -0.062$, $df = 399$, $p = 0.212$). No significant correlation was found between the mean of the priority levels and the mean severity ratings ($r = -0.08$, $N = 400$, $p = 0.112$).

Considering WCAG 1.0, 330 distinct website were covered by WCAG 1.0. No significant correlation was found between highest priority levels of related CPs and the mean severity ratings of user problems ($r = -0.09$, $N = 330$, $p = 0.102$). No significant correlation was found between mean priority levels of related CPs and the mean severity ratings of user problems ($r = -0.039$, $N = 330$, $p = 0.484$).

5.6.2 Severity of problems encountered by partially sighted users and priority of related WCAG checkpoints/success criteria

For partially sighted users, 259 distinct website problems were covered by WCAG 2.0. No significant correlation was found between the highest priority levels of related SCs and the mean severity ratings of user problems ($r = -0.021$, $N = 259$, $p = 0.741$). No significant correlation was found between the mean of the priority levels and the mean severity ratings of website problems ($r = -0.039$, $N = 259$, $p = 0.535$).

Considering WCAG 1.0, 195 distinct website were covered by WCAG 1.0. A significant, but low correlation was found between the highest priority level of CPs related to website problems and the mean severity ratings of website problems ($r = -0.175$, $N = 195$, $p = 0.027$). A significant, but also low correlation was found between the mean priority level of CPs related to website problems and the mean severity ratings of those ($r = -0.186$, $N = 195$, $p = 0.009$).

5.6.3 Severity of problems encountered by dyslexic users and priority of related WCAG checkpoints/success criteria

For dyslexic users, 85 distinct website problems were covered by WCAG 2.0. No significant correlation was found between the highest priority levels of related SCs and the mean severity ratings of user problems (r = -0.017, N = 85, p = 0.876). No significant correlation was found between the mean of the priority levels and the mean severity ratings (r = -0.058, N = 85, p = 0.599).

Considering WCAG 1.0, 39 distinct website were covered by WCAG 1.0. No significant correlation was found between highest priority levels of related CPs and the mean severity ratings of user problems (r = -0.024, N = 39, p = 0.885). No significant correlation was found between mean priority levels of related CPs and the mean severity ratings of user problems (r = -0.023, N = 39, p = 0.892).

5.7 Summary of the chapter

This chapter presented the results and discussions related to the secondary research question proposed in this thesis, regarding the relationship between problems encountered by print-disabled users on websites and technical web accessibility guidelines. The chapters presented results showing differences between the number of problems encountered by print-disabled users in websites that conformed and did not conform to WCAG 1.0 and 2.0. Following, the chapter presented analyses on the correlation between the number of users problems in websites and measures of the number of instances of violations of checkpoints/success criteria and the number of different checkpoints/success criteria violated.

The results in the chapter also presented an analysis of the percentage of problems encountered by print-disabled users that are covered or not by WCAG 1.0 and 2.0, followed by details of the types of problems encountered by each user group that are not covered.

For problems that were covered by the guidelines, the chapter presented an analysis as to whether the web pages where problems occurred had implemented relevant guidelines successfully or not. In cases where relevant guidelines had been implemented, analyses were performed to identify the types of problems related to those guidelines and why guidelines failed to prevent them.

Finally, the chapter presented an analysis that revealed a lack of significant correlations between the severity ratings of user problems and the priority levels associated with checkpoints/success criteria.

The next chapter presents a general discussion of the results obtained in this thesis and how they addressed the research questions proposed.

Chapter 6. Discussion

This chapter presents a general discussion of the findings of the study reported in this thesis. The chapter relates the outcomes of the study to each of the main research questions.

The study had as the main research question to investigate *what are the main characteristics of accessibility problems encountered by print-disabled users when attempting to use websites*, and a secondary research question to *investigate the relationship between user-based measures of accessibility of websites and measures of technical web accessibility based on the guidelines defined in WCAG 1.0 and 2.0*.

In order to answer those research questions accessibility evaluations of a range of websites were performed with a sample of 64 disabled users – particularly those with print disabilities, being 32 blind, 19 partially sighted and 13 dyslexics. The sample of websites evaluated had 16 websites, including websites at different levels of conformance to WCAG 1.0 and WCGA 2.0, in order to enable further analysis of the relationship between user problems and technical guidelines and address the secondary research question.

The evaluations with print-disabled users yielded 3,012 problems encountered by users, to which severity ratings were assigned. Those problems were classified into categories in order to better understand the types of problems encountered by different user groups. The analysis of those problems was the core resource to address the main research question, as described in Section 6.1 of this chapter, along with the analysis of task completion and difficulty to complete tasks.

After the identification of problems encountered by users, those problems were compared to technical web accessibility guidelines in WCAG 1.0 and WCAG 2.0, in order to analyse the coverage of those problems by the guidelines and analyse the relationship between user problems and measures of technical web accessibility. Those analyses were performed to address the secondary research question, as described in Section 6.2 of this chapter.

Section 6.3 presents limitations of the research, and Section 6.4 presents a summary of this chapter.

6.1 The characterisation of accessibility problems encountered by print-disabled users

The primary research question proposed in the research reported in this thesis had the aim to characterise the main types of accessibility problems and measure the accessibility of websites in terms of whether print-disabled users can use them or not.

The results provided a very substantial body of evidence with 3,012 problems encountered by a panel of 32 blind, 19 partially sighted and 13 dyslexic users, which makes this one of the largest studies on the accessibility of websites involving print-disabled users. In order to address the question, several sub-questions were proposed to investigate different aspects of the use of websites by print-disabled users, including whether they can complete the tasks they attempt, the difficulty to perform tasks, and, most importantly, the main types of accessibility problems they encounter, as well as their frequency and severity.

The following sections present a general discussion of how those sub-questions were addressed in this thesis and the main findings related to them.

6.1.1 Task completion on websites by print-disabled users

The present study revealed important findings regarding the task success rates on websites by print-disabled users. The results presented in Section 4.1.1 showed that print-disabled users still have problems to complete tasks on websites.

Blind and partially sighted users had lower task success rates than dyslexic users, and could not complete more than 40% of the tasks they attempted. Blind users could only succeed in 56% of tasks attempted and partially sighted succeeded in only 49%, while dyslexic users succeeded in 84.96% of the tasks. This indicates that accessibility problems in websites create severe barriers to those users and prevent them from completing every-day tasks on websites, in particular to blind and partially sighted users.

In the study performed by the Disability Rights Commission (2004), blind users succeeded in 53% of tasks attempted, partially sighted users in 76% and dyslexic users in 83%. The results of task success rates for blind and dyslexic users in the present study were very close to the results obtained in the DRC study. However, partially sighted users succeeded in substantially fewer tasks than in the DRC study.

These results show that, like in 2003 (when the DRC study was conducted), disabled users still encounter a lot of problems in every-day tasks they attempt to complete on the Web. The results in the present study were particularly alarming for blind and partially sighted users, who had problems to complete more than 40% of the tasks they attempted.

It is also very alarming that between 2003 and 2010 (when most of the evaluations in the present study were conducted), the number of tasks that blind and partially sighted users could complete did not have any improvements, or what is of greater concern, became even worse. As a measure of users being able to do tasks they want to on websites, this result shows that accessibility has not improved since then.

The lower completion rate for partially sighted users in comparison to blind users was not expected from previous results in the literature. In the DRC study, for example, partially sighted users had a higher task success rate than blind users. One of the possible causes might have been the lack of matching of computer experience between users from different groups. The level of computer experience of blind users in the present study was higher than that of partially sighted users. Another possible cause may be the different sight conditions of users in the two studies. Many partially sighted users in the present study had very severe sight loss and needed to use very high levels of magnification, which may have had a more severe impact on their use of websites. The use of more complex visual layout structures and new technology may also be a cause of the lower success rates for partially sighted users, as websites have gone through many changes between 2003 and 2010. This is supported by the large number of problems encountered by users in the present study that were related to multimedia (especially with embedded videos) and highly interactive interface components (more dynamic content and interactive applications, such as those using Flash), that have become much ubiquitous in websites than they were in 2003.

The unexpected results from partially sighted users show that more attention needs to be given to this group, as they can be severely affected by problems in websites which prevent them from succeeding in their tasks. It has been acknowledged that blind users can have more difficulties to complete tasks due to particularities of the use of screen readers, which leads to developers having to think carefully about how to design the interaction for those users. However, the results in this study showed that the impact of the problems partially sighted users can encounter due to the ways in which they interact with websites may have been underestimated, and can have a comparable impact to that of blind users. When users have a limited viewport to see only a small portion of the screen and have to pan across the screen to see all content, for example, users can spend a substantial amount of time looking for content and can potentially miss crucial information and functionality that is on the screen to complete their tasks. In comparison with blind users, partially sighted users encountered some problems in common, but other problems were very distinct.

6.1.2 Difficulty of performing tasks on websites by print-disabled users

The findings in the present study about the difficult of performing tasks by print-disabled users were very important in identifying how difficult different user groups find to use websites. The results are presented in Section 4.1.2.

In agreement with the findings about task completion by different groups, blind and partially sighted users also had the highest difficulty ratings when compared to dyslexic users. As with task completion rates, partially sighted users in the present study also had the highest ratings of difficulty to perform tasks on websites, when compared to those of blind users and dyslexic users.

In the DRC study, a different ratings scale was used, with 7 meaning “very easy”. In that study, the mean ease of task rating for blind users was 4.2, 5.1 for partially sighted users and 5.6 for dyslexic users. Although the present study found similar results with dyslexic users finding tasks less difficult than blind and partially sighted users, the difficulty encountered by partially sighted users in tasks in this study was comparatively much higher than in the DRC study.

The absence of statistical significance between task difficulty ratings from blind and partially sighted users is also an interesting result. Along with the results from task success rates, it also suggests that partially sighted users may also encounter comparable difficulty in performing tasks as blind users do.

Although blind and partially sighted users are very different in the way they interact, the results in this study showed that they shared many problems that impacted on the difficulty they had to use websites. Both blind and partially sighted users had severe problems when trying to navigate websites to find content and navigation structures were unhelpful. This can cause serious difficulties to those users when they have to understand how a website works given the limitations of their assistive technologies or adaptations.

Blind and partially sighted users have to split their cognitive effort between trying to build an overview of the website from limited chunks of information (linear reading of content for blind users, or show small portions of the screen at a time for partially sighted users) while also dealing with usual usability problems that also affect mainstream users. This extra cognitive effort can make performing tasks on websites for those users much more difficult than for other users.

6.1.3 Instances of problems per user group

Blind and partially sighted users encountered considerably more problems per website than dyslexic users, with a mean number of instances per website per user of 9.22 for blind users, 8.09 for partially sighted users and 4.64 for dyslexic users. This reinforces the notion that blind and partially sighted users were more affected by accessibility problems than dyslexic users.

The absence of statistical significance between the mean number of problems encountered by blind and partially sighted users shows that partially sighted users encounter a similar amount of accessibility problems when using websites, which might be contrary to common belief that partially sighted users do not have as many problems as blind users.

The amount of problems with inadequate presentation of graphical elements on the screen makes partially sighted users encounter problems very frequently. This, along with other frequent problems, such as those related to unhelpful navigation structures, make for a high number of user problems for partially sighted users. This shows that designers still fail to make interfaces with presentation that works well for partially sighted users and navigation structures that helps them find content quickly and minimises the burden of searching for information on websites. However, it is important to highlight that, given the high frequency and severity of those issues, making good efforts to design more adaptable presentations and improving navigation structures in websites alone would bring a very significant reduction in the number and impact of accessibility problems encountered by partially sighted users.

The similarity in the higher number of problems encountered by blind and partially sighted users on websites is in line with the findings in this study that they are also the user groups that have the most problems to succeed in their tasks and the highest difficulty levels to perform tasks.

6.1.4 Main types of problems encountered by print-disabled users on websites, their frequency and severity

While the sub-questions presented previously yielded important findings about how print-disabled users performed tasks on websites and the number of problems they encountered, the key sub-questions related to the primary research question proposed in this study were questions regarding the characterisation of the types of problems encountered by print-disabled users, the frequency that they occurred and their severity level.

The following sections present the discussions of problems found by blind, partially sighted and dyslexic users, respectively, followed by a description of some of the main problems that were shared between user groups.

6.1.4.1 Problems encountered by blind users

Blind users encountered a wide range of accessibility problems, which is shown by the larger number of subcategories covered by their problems. Many problems were particular to blind users, and other problems were also common to partially sighted and dyslexic users, but affected users in different ways.

Many of the types of problems encountered by blind users in the present study were also encountered by previous large studies on the accessibility of websites, such as the study performed by the Disability Rights Commission (2004) and the study performed by Coyne and Nielsen (2001). Problems in common with other studies included issues like: incorrect or non-existent labelling of links and form elements, cluttered and complex page structures, ALT tags on images non-existent or unhelpful, confusing and disorienting navigation mechanisms, issues with pop-up windows, problems with links and buttons and tables. Unlike the study performed by Coyne and Nielsen (2001), the present study did not include many problems related to frames. This is probably due to a reduction in the use of these elements in the layout of web pages.

Although some issues were common with previous studies, the findings from this study provided very important insights into problems encountered by blind users that can help understand how blind users use websites, to understand the nature of the problems they encounter, how those problems impact blind users, and actions that can be taken by web developers to avoid those problems.

The first step to better understanding what problems matter the most to blind users is looking at the top-5 most critical problems in terms of severity and frequency.

The first type of problem listed as top-5 most critical was the inability to reach controls or form elements using the keyboard. Making sure interfaces can work appropriately with a keyboard is an essential aspect to make websites accessible to blind users. The use of more interactive elements in web pages can be very dangerous to the accessibility to blind users if developers do not take special care to make sure all functionality works well for users who only use a keyboard, and not a pointing device such as a mouse. The rapid growth of Web 2.0 applications is likely to increase significantly the number of such interactive elements on websites, which could make this type of problem potentially more prevalent.

The significant growth in the use of multimedia and videos on the Web brought a very important issue to blind users to the spotlight in the top-5 most critical problems: providing audio descriptions of videos. Many blind users expressed frustration when they could not fully experience video content on websites because they did not have descriptions in audio of non-speech content in videos, such as the description of scenes and actions. Blind users want to use videos on the Web more and more, and designers and content creators need to include creating audio descriptions of videos in their priority lists if they want blind users to have access to those videos in the same way as their audience of mainstream users. It is worth noting that some guidelines give the option of providing text transcripts with full descriptions of scenes as an alternative to audio description. However, as discussed in Section 6.2.5, this type of implementation was rejected by blind users.

Problems with unhelpful navigation structures were also listed as the most critical, especially due to its high frequency. Although this problem was shared by other user groups, it is worth noting how this problem impacts blind users. Exploring different options in the navigation by trial and error can be substantially difficult to blind users due to the nature of the assistive technologies they normally use. Having to go back and forth in the navigation of a website can make users become lost, besides adding to the cognitive effort that blind users have to make to manipulate their screen readers and understand the structure of a website. Making navigation bars that indicate quickly where to go can have a substantial impact on blind users, considering the time and effort they can save.

Problems with controls and form elements that do not have a clear description of their purpose were considered critical due to their very high severity.. Designers often forget about providing descriptions of controls that are accessible to blind users via their screen readers, such as with labels and textual descriptions. For blind users, it is essential that a description is provided in text that is accessible to screen readers on a web page's source code. The concern for providing accessible descriptions should be observed not only in regular HTML pages, but also in embedded applications, such as those using Flash.

Problems with no or insufficient feedback were also very critical to blind users. Besides having a concern for providing good and informative feedback messages to users, developers also need to think about whether blind users will receive their messages or not. Many websites include feedback messages added dynamically to pages, and blind users were not informed of this update, making it difficult to locate. If feedback messages are not clear and well positioned in a page to help blind users find them (e.g. in a place that can be found with a heading or in the beginning of the page), they can potentially go "unheard" by them, seriously jeopardising the interaction of those users.

Besides the top-5 most critical problems encountered by blind users, this study also revealed other very important findings about blind users and websites.

Compared to older studies, the present study showed different results in relation to the use of different interactive technologies and multimedia, especially with Flash. Unlike older studies that suggested that a completely separate alternative to applications in Flash was necessary, this study showed that many blind users can now use basic features of Flash applications if basic accessibility requirements are implemented. The improvement in the support of those technologies is probably the reason for the differences. However, whilst in older studies the problems blind users were that they were not able to access interactive content and multimedia at all, the present study revealed many accessibility problems are still encountered by blind users within those applications. A lot more needs to be done to ensure that blind users cannot only have access to basic features in Flash applications, but that they can use them satisfactorily, covering other issues such as having appropriate feedback, ensuring effective navigation within Flash applications and providing good indications of how to interact with interface components. Despite the increase in the support of those technologies by screen readers, it is still very important to take care to implement accessibility features in their interactive applications to make them accessible.

Another very interesting finding from this study was the extensive use of headings by blind users to navigate within pages to find information. A survey conducted by

WebAIM (2011) with of 1,245 screen-reader users showed that 57.2% of the users surveyed use headings as a first strategy to find content in a lengthy page. The types of problems blind users encountered in this study provide valuable information for developers to avoid problems related to the misuse of headings and to make a heading structure that optimises the navigation in a web page by blind users. For users who prefer using headings as a primary strategy to explore the content in a web page, they can become disappointed if no headings are available on pages and they cannot use their first choice of information finding strategy. Designers should also make sure all headings are properly marked-up as such, so blind users can find them easily using their screen readers. Special attention should also be given to making heading structures that provide a good overview of the page, conveying how the topics are organised in a page. Another interesting fact was that nearly 100% of problems related to headings were specific to blind users. Although some problems encountered by other user groups could be addressed with headings, such as difficulties to scan for specific items, but users did not complain specifically about headings in those cases. In problems with headings for blind users, users attributed problems directly to headings, and gave very specific reasons as to why specific issues with headings (of lack of them) were a problem to them. This shows that the navigation in web pages by headings is a strategy that has been widely adopted by blind users as a particular strategy of this user group. However, after blind users having adopted headings so widely and depending so much on them, this implies that not having good headings that are properly marked-up can be especially severe to blind users.

The severity rating of problems related to textual descriptions of images was another interesting finding. This study showed that, unless an image conveys essential information to users, blind users considered problems with lack of or inappropriate textual descriptions as annoyances, but not as severe problems. This does not mean, however the provision of textual descriptions to non-essential images should be abandoned. Although the lack of descriptions for one or other image can be just an annoyance, when many of those problems mount in a web page, blind users can be serious jeopardised in their navigation. Inappropriate textual descriptions, like those with file names or codes, means that a lot of non-sense content is read to users, who have to spend more time trying to skip that content and get to the content they want.

Unlike improvements in Flash, that seemed to have been better accepted by users, there was a very strong aversion to text in PDF format by blind users. Although some developments have been made to incorporate accessibility features into PDF documents, it seems like those did not have a significant effect on users at the time the study was conducted. Even when PDFs were properly marked-up, blind users felt they did not have the same flexibility and features to read as in other file formats. Some users with more aversion did not even want to try to use PDFs initially. However, even users who tried to use them expressed dissatisfaction when they could not use their usual reading and navigation strategies in PDF documents as they could in other types of documents. Issues with PDF also affected partially sighted and dyslexic users, in that there was greater difficulty to use PDFs with screen readers, and many dyslexic users could not change colour background and font settings on PDFs.

The way blind users read links was another important finding in this study. A substantial amount of blind users use features in their screen readers to list only the links in a web page. The amount of problems related to users not being able to identify the destination of a link showed it is a very frequent issue, and one that should be carefully observed by designers. They should be aware that a substantial amount of blind users will read their link texts out of a context in a list with only links. If the link text is not meaningful on its own, this can potentially result in blind users not being able to know where they will take them.

The use of security checks with CAPTCHAS remains as an issue to be solved. This study showed that providing an audio CAPTCHA with excessive noise to the recognition of distorted text in an image is not adequate to blind users. Trying to decipher letters spoken on top of noisy sounds was very challenging to blind users, and took them a considerable time to even try to solve. If audio alternatives are to be provided, they need to be carefully designed and tested to verify if it is really possible for blind users to understand them. For example, Lazar et al. (2012) proposed a different type of audio CAPTCHA in which users have to recognise sounds such as bells or a piano, and tested the proposal with blind users. Their study showed that there was a significant improvement in the success rates in decoding those CAPTCHAS. However, the report states that more tests need to be performed with real applications to verify how the system would work.

6.1.4.2 Problems encountered by partially sighted users

The problems encountered by partially sighted users in the present study also covered a wide range of types of problems, with the number of different subcategories of problems second only to blind users. A total of 18 subcategories that had the majority of problems encountered exclusively by partially sighted users. Although these represent fewer subcategories than those encountered exclusively by blind users, this is a very representative set of problems and shows the importance of involving partially sighted users in the evaluation of websites.

Many of the types of problems encountered by partially sighted users in the present study were also encountered by previous large studies on the accessibility of websites, such as the study performed by the Disability Rights Commission (2004) and the study performed by Coyne and Nielsen (2001). Problems in common with other studies include issues like: inappropriate use of colours and poor contrast between content and background, incompatibility between accessibility software (e.g. for magnification) and web pages, unclear and confusing layout of pages, confusing and disorienting navigation mechanisms and graphics and text size too small.

As in the case with blind users, the results from partially sighted users provided very important insights to understand how they use websites, the problems they encounter and how to avoid those problems.

Of the top-5 most critical problems encountered by partially sighted users, listed according to their severity and frequency, three were related to problems with the presentation of graphical elements – namely controls/form elements, images and text. Many of those problems were related to problems with colour, size and resolution. In particular, more than 90% of problems related to the presentation of controls/form elements and images were particular to partially sighted users and not shared by other user groups.

Problems with controls/form elements were particularly severe to blind users. Those elements are normally associated with features available on websites. If partially sighted users have problems reading them, this means that they will probably have serious problems to identify what they need to do to complete their tasks.

Testing how a web page is rendered to partially sighted users can be very challenging to developers. They can use a range of different assistive technologies and adaptations in their operating systems and browsers, that can involve magnification or different screen resolutions, different colour schemes (such as high contrast, inverted colours, black/white, for example). However, it can be very challenging for designers to design a single website that would accommodate all possible different settings that different user profiles might need. Designing websites that could accommodate personal adaptations according to different user profiles would be a more promising solution, as discussed by Theofanos and Redish (2005). Although not all issues encountered by partially sighted users are well accommodated by existing technology, it is important that designers consider some issues that can be addressed at the moment.

In the top-5 most critical problems encountered by partially sighted users were also two types of problems related to unhelpful navigation and users not finding content in places where they expected. Although common to other user groups, those problems also impact on partially sighted users in different ways. Partially sighted users can have problems with navigation structures that are aggravated when they can only see part of the screen, making them spend a considerable amount of time to browse through different possible options. As for blind users, designing good navigation structures also can have a significant bonus to partially sighted users in reducing the time they spend exploring different navigation options and juggling with their assistive technologies or dealing with pages that have their sizes significantly increased by their web browsers.

Like issues with the organisation of information to navigate between pages in websites, organising information within web pages was also very important to partially sighted users. Logically organised pages meant that users who only see a small viewport could find more easily the content they wanted instead of having to roam around a page to find a piece of information.

Opening links in windows is a well-known issue to blind users, as they can become lost by not knowing which windows are open to keep track of. Findings from this study showed that avoiding opening new windows without users' knowledge is also an important issue to partially sighted users as well. Especially with users of screen magnifiers, they cannot always see everything in the screen, and may not recognise that a new window was opened. Those findings show that partially sighted users too can benefit from being told about new windows opening when clicking on a link, avoiding them feeling lost and potentially closing a window by accident later and losing work they had done. Current implementations that inform this to users normally use icons in front of a link to identify that it opens in a new window, with the icon having an alternative text normally identifying "opens new windows". While having the alternative text in the alternative text to an icon makes it be read to screen readers, it is not clear if such approach would be effective to partially sighted users. Many users might not be able to recognise what the icon means to them, and depending on how the icon is designed, they might not even be able to see it. More specific studies with partially sighted users are necessary to determine what strategies work best to inform those users that links open new windows.

The use of multimedia, especially videos, on websites also presents some challenges to partially sighted users. As well as presentation of images, text, and controls/form elements, many users encountered problems with the presentation of videos or animations, such as problems with size, colour contrast and speed of presentation of videos, or with videos that simply were not compatible with screen magnifiers. Those problem types were particular to partially sighted users. This shows that content producers should test too how multimedia and video content is displayed using different colour and size settings, and testing with screen magnifiers. In the case of multimedia animations, tests can be performed more easily to determine if it can cause problems for partially sighted users to visualise. For videos, however, this can be a much more challenging issue, as videos can be of filming of scenes in natural environments, where it would be difficult to control such issues.

An important remark is related to the high severity ratings of problems encountered by partially sighted users in comparison to the severity ratings from blind and dyslexic users. This result reinforces the importance of looking at partially sighted users as a critical user group when dealing with web accessibility, who are severely affected by accessibility problems.

6.1.4.3 Problems encountered by dyslexic users

The results from the evaluation of websites by dyslexic users in this study provide important contributions to advancing the body of empirical evidence of accessibility problems encountered by those users, which has received far less attention than other user groups.

The present results are consistent with results from other studies with dyslexic users (Al-Wabil et al. 2007, Bradford 2005, British Dyslexia Association 2011, Disability Rights Commission 2004, Evett and Brown 2005, Rello et al. 2012, Zarach 2002), especially those regarding layout-related issues, such as problems with typeface, colour and text disposition, issues related to difficult language, and others related to navigation (Al-Wabil et al. 2007, Disability Rights Commission 2004). The present study extended previous results by providing insight into the context in which dyslexic encounter several types of problems.

The study presents, though, very important information that can provide clues as to how to implement solutions to many of those issues in websites, and also included other problems that are more related to website issues.

In the top-5 most critical problems encountered by dyslexic users were problems with unhelpful navigation structures and users not finding content where they expected. The first type of problem was also in the top-5 critical to blind users and partially sighted users, and the second in the top-5 critical to partially sighted users. However, as for blind and partially sighted users, those problems also have particularities that affect dyslexic users in different ways. Some dyslexic users claimed that an important part of understanding a website for them was "having a sense of structure". If navigation structures are not helpful, or users do not find information where they believed was the right place, their confidence on the mental model of the website they formed can erode, and this can make users get lost in websites and find it harder to browse through it.

Users expecting a certain functionality not present was also in the top-5 most critical problems encountered by dyslexic users. The feature that most users wanted to have was the auto-complete feature to help with spelling. Many users became used to using this in search features, and stated they find it very beneficial to have it in form fields in websites, so they can have help to spell words they find difficult. This is an indication for developers to investigate the possibility of implementing this feature in their websites whenever possible, with the possibility of using services from other websites. Another feature that was requested by users was a search feature. Many users stated that their first approach to finding information in a website is use the search feature. When this feature was not present, this represented a problem as users could not use their first information finding strategy. This suggests that developers should consider including internal searches in their websites, to enable users to find information more quickly.

Problems in which users could not make sense of content were also in the top-5 most critical encountered by dyslexic users. Users find it difficult to make sense of content when websites presented incomplete information or abbreviations users did not understand. It is very important that content producers pay special attention to making sure that content is written clearly, and that users can make sense of the messages conveyed without difficulties. Incomplete information that is scattered across different pages, or abbreviations with definitions outside of a page may make it very difficult for users to understand what is contained in websites.

The last problem type in the top-5 most critical encountered by dyslexic users was problems with malfunctioning features. It was very interesting that a significant amount of those problems were related to search features that did not find information that users expected was on the website. Following the finding that many dyslexic users want to have a search feature in websites, this showed that they also become disappointed when the search does not return expected results. Developers should test carefully the implementation of search features and how they are integrated into websites. A common issue that caused searches not to work was that content in static pages were not indexed, for example.

Besides the top-5 most critical problems, other findings of other types of problems encountered by dyslexic users were very interesting. Many problems related to the presentation of text were encountered, including problems with font typeface, size, colour, alignment and use of columns. As stated by existing guidelines, it is recommended to avoid the use of Serif font, text that is too small, justified paragraphs, parallel columns, italics and black writing on white background. With regards to the issue with colour background and foreground, an interesting finding was that most users that needed this adaptation expected that websites would provide them with an embedded feature. This is an important information for designers, so they can be aware of what users expect, and most importantly, know that many users will most likely not use the colour settings on web browsers. That means that, for those users, if such a feature is not available, dyslexic users will continue having difficulties to read text with black writing on white background on their websites.

Dyslexic users also had problems where they found it difficult to scan for content on web pages. Providing good visual cues to help users scan for important topics and important information can help greatly users find information more quickly. Designers should consider marking properly titles, headings, sections and special keywords to help users find them more quickly.

Another interesting issue was that a significant amount of website problems with navigation bars not being salient were identified exclusively by dyslexic users. This shows that dyslexic users can be particularly affected by issues of having information that does not stand out, especially when it is a crucial element such as the navigation bar. Designers should pay special attention to design navigation bars that are positioned in prominent places in web pages and properly formatted to make it easier for users to find them.

6.1.4.4 Common problems encountered by different user groups

Two problem types had a substantial number of website problems that were encountered by all user groups in this study: “*navigation elements do not help user find what they are seeking*” and “*content not found in pages where expected*”. Although those problems can affect several different groups of users with disabilities, the discussions in Section 6.1.4.1, Section 6.1.4.2 and Section 6.1.4.3 showed that different groups of users with disabilities can be affected differently by those problems. For blind and partially sighted users, for example, unhelpful navigation elements may become a heavy burden given that they have the cognitive effort of dealing with their screen readers and understanding the website. For them, having to explore several different links can also be a problem due to the effort needed to go back and forth in different pages. The same happens to partially sighted users, as they have to spend a lot of effort to go back and forth in pages to find the right link to a content, especially if they use a screen magnifier with a small viewport. For dyslexic users, more structured navigation structures can help improve the confidence in websites and reduce the effort to understand the organisation of websites.

Another issue that was common to the list of most frequent problems of several user groups was not having an expected feature, especially a search in websites. It seems using search to find information has become a norm to users, and they become disappointed if websites do not offer them this feature and they have to browse through all the navigation.

Besides expecting to have a search feature, all user groups encountered problems with search features that do not work. If a search feature is available, users from all groups became disappointed if some information was not found by it.

Not being able to return the home page was also an issue encountered by users from all groups. It is very important that users are able to locate a link that takes them back to the home page of a website any time they want, as described by Nielsen's user control and freedom heuristic. However, in some cases problems may be related to other issues related to the organisation of websites. For example, in many websites evaluated in this study, developers create "sub-sites" that are linked from the main site. A serious problems with those sites was that they often have some identification that they were part of the main site. However, their navigation structure was separate from the main site, and links that were named "Home" linked to the home of the sub-site, not the main site. When developers choose to create sub-sites, it is very important to keep in mind that some users might have arrived from another related website and that they want the main principles they know about the website to keep working, for example, offer an option to return to the main home page where they started their navigation from.

6.2 The relationship between user-based measures of accessibility of websites and measures of technical web accessibility based on WCAG 1.0 and WCAG 2.0

This thesis presented important findings resulting from the analysis of the relationship between user-based measures of accessibility of websites and measures of technical accessibility guidelines. The secondary research question was addressed by six sub-questions that involved comparisons and analyses of different aspects related to the conformance of websites to WCAG 1.0 and WCAG 2.0.

This section is organised to present detailed discussions related to each of the six sub-questions covering particular aspects of the relationship between problems encountered by disabled users and measures of technical web accessibility with WCAG 1.0 and WCAG 2.0.

6.2.1 The relationship between number of problems encountered by print-disabled users on websites and conformance to WCAG 1.0 and WCAG 2.0

The relationship between the number of problems encountered by print-disabled users in websites at different levels of conformance to WCAG was the first aspect analysed regarding the secondary research question proposed in this thesis. The results of the analyses are presented in Section 5.1.

The analysis of the first aspect of the technical guidelines in WCAG 1.0 and 2.0 yielded important findings for web developers. The findings about the difference in user problems in websites at different levels of conformance in this study are a very important contribution in relation to current knowledge in the field. Previous related studies (Disability Rights Commission 2004, Rømen and Svanæs 2008, Rømen and Svanæs 2011) did not include in the study design websites at difference conformance levels to enable to perform the analyses presented in this thesis, which makes those findings a novel and important result.

Despite the importance given to making websites that achieve certain levels of conformance, the results showed that higher levels of conformance may not translate into print-disabled users encountering fewer problems on websites. A number of factors that were investigated in this thesis can be possible causes of this. The results obtained showed that a substantial amount of problems encountered by disabled users were not covered by the checkpoints (CPs) /success criteria (SCs), being adherence to them one of the main aspects considered to establish conformance. The results also showed the lack of correlations between measures of conformance, such as the number of instances of violations and number of different violations of checkpoints/success criteria (except for number of different CPs/SCs violated for blind users) and problems encountered by print-disabled users. Another cause could be that the implementation of the requirements of many problems covered by guidelines does not necessarily avoid user problems. With all those measures that are strongly related to conformance are not related to user problems, it is no surprise that increasing conformance levels does not necessarily reduce the number of problems encountered by disabled users.

When comparing results from different user groups, blind users were the only group that had a decrease in the numbers of user problems in websites that were not conformant to WCAG 1.0 to websites at WCAG 1.0 level A, and from non conformant to WCAG 1.0 level AA. However, it was very surprising that there was no significant decrease between level A and level AA. Achieving level AA may mean significant effort from companies to implement all the requirements. It is very concerning that all this effort does not necessarily translate into print-disabled users encountering fewer problems, as would be the likely goal of websites that strive to achieve those levels.

For WCAG 2.0, it was not possible to perform the same tests comparing websites at all levels individually, as there were so few websites that were conformant. However, when comparing the number of problems encountered by blind users in websites that were non-conformant with websites at level A, and non-conformant with websites conformant at any levels, neither comparison showed any significant decrease in the number of problems.

This finding seems to point to suggest that the upgrade to WCAG 2.0 did not have the effect that was expected. Given that the new version of the guidelines took nearly ten years to be published since the first version, one would expected that improvements would be made that would be directly reflected on the reduction of the number of problems that print-disabled users encounter on websites that conform to WCAG 2.0.

For WCAG 2.0, one would expect there to be a larger decrease in the number of problems encountered by blind users from non-conformant websites to Level A conformant websites than there was for WCAG 1.0. However, the results showed that conformance of a website to WCAG 2.0 Level A does not mean that users will encounter fewer problems on it and as a result it does not necessarily mean that following WCAG 2.0 will “make content accessible to a wider range of people with disabilities” (Caldwell et al. 2008).

When the same analyses performed for blind users were made for the number of problems encountered by partially sighted and dyslexic users, an even worse scenario was found. No reductions on the numbers of user problems were found when comparing non-conformant with websites that were conformant at any levels with either WCAG 1.0 and WCAG 2.0. This shows that both WCAG 1.0 and WCAG 2.0 conformance levels do not have any impact on those user groups.

It has been argued that support by WCAG 1.0 and WCAG 2.0 for users with cognitive disabilities and learning difficulties is lacking Seeman (2006). The findings in this study supports this argument, and shows that conformance to either version of WCAG fails to reduce the number of problems encountered by dyslexic users.

Partially sighted users are one of the groups that seems to receive significant attention in motivational and introductory accessibility descriptions in the guidelines (Cooper et al. 2010a), and one would expect that a higher effect of conformance would obtained for those users. However, the findings in the present study show that conformance to WCAG 1.0 and WCAG 2.0 does not mean that partially sighted users will encounter fewer problems on websites.

Those results showed that achieving certain conformance levels with WCAG 1.0 and WCAG 2.0 can be very ineffective as a means to reduce the numbers of problems encountered by disabled users. The way the conformance requirements are structured do not seem to address the all-important concern of making websites that disabled users can use better and encountering fewer problems.

6.2.2 The relationship between the number of instances of problems encountered by print-disabled users on websites and violations of WCAG 1.0 checkpoints/ WCAG 2.0 success criteria

Besides the overall conformance levels of web pages, the second aspects regarding the relationship between user problems and the technical guidelines was the number of violations and the number of different checkpoints/success criteria violated. The results from the analyses are presented in Section 5.2.

Two measures related to the violation of WCAG CPs/SCs were used: the number of *instances of violations* and the number of *different CPs/SCs violated*. With regards to the number of instances of violations, no correlation between the mean number of user problems per website per user was found for any of the user groups for either version of WCAG. This means that the number of times that a given website violates WCAG CPs/SCs is not a good predictor of how many problems print-disabled users will encounter on those websites. It is important to be noted that many types of violations can be related to purely technical issues that do not necessarily cause direct problems to users. For example, in WCAG 2.0, each HTML mark-up error is one violation of SC 4.1.1, and, depending on the type of violation, it may not create a barrier to users.

Regarding the number of different CPs/SCs violated, no significant correlation was found for the number of problems encountered by partially sighted or dyslexic users. However, significant correlations were found between this measure and the number of problems encountered by blind users for both WCAG 1.0 and 2.0. This could be possibly explained by the fact that a substantial number of WCAG 1.0 CPs and WCAG 2.0 SCs relate to issues that are more connected to problems encountered by blind users than by other disability groups. It is interesting that this correlation was found, when there was no correlation between the number of user problems and the number of instances of violations. A possible explanation to this correlation could be that when a website violates fewer of those CPs/SCs, it could mean that it is less likely to present a combination of different types of barriers that can create more problems to users than a number of problems of the same type. Given that a significant number of CPs/SCs are related mostly to blind users, violating fewer of those could mean a greater concern from the website with blind users.

6.2.3 The coverage of problems encountered by print-disabled users on websites by WCAG 1.0 and WCAG 2.0

One of the key issues to address the secondary research question proposed in this thesis was to establish whether problems encountered by print-disabled users on websites were covered by WCAG 1.0 and WCAG 2.0, and when covered, if relevant guidelines were implemented or not in web pages where problems occurred. The results that addressed this sub-question were reported in Section 5.3.

The analysis of the coverage of problems by WCAG 1.0 and WCAG 2.0 revealed interesting findings in the present study. The amount of problems that were covered for blind, partially sighted and dyslexic users provide a very good estimate of the extent to which guidelines can help them uncover accessibility problems in websites, and how much needs to be addressed by other methods, in particular with evaluation by disabled users. For the problems that were covered by guidelines, this study also analysed whether web pages where problems were found successfully implemented the requirements for WCAG. The results from this analysis are very relevant to establish how effective guidelines are to prevent user problems.

It was very surprising that, even for blind users, the percentage of problems covered was only at around 50%, and even lower for partially sighted and dyslexic users. However, the percentage of problems. The total coverage for all user problems by WCAG 1.0 was around 29% of problems only, and 38% by WCAG 2.0.

The percentage of problems covered by WCAG 1.0 in the present study is even lower than the percentage found in the study performed by the Disability Rights Commission of Great Britain (2004), that found that around 45% of problems were not covered by WCAG 1.0. However, the results in the DRC study also included users with hearing and physical impairments, which could have had different coverage patterns. The percentages of problems covered by guidelines in this study were similar to those encountered in the study performed by Rømen & Svanæs (2008, 2011), in which 27% of problems encountered by blind, partially sighted, physically impaired and dyslexic users were covered by WCAG 1.0 and 32% were covered by WCAG 2.0.

By analysing the criteria of coverage and implementation of guidelines, the total set of user problems can be divided as described in Figure 6.1. In this figure, the outer layer represents user problems that are not covered by guidelines, the middle layer represents problems that are covered, but guidelines were not implemented in web pages, and the innermost set, with problems that were covered by guidelines, and had web pages that successfully implement them, but were not effective.

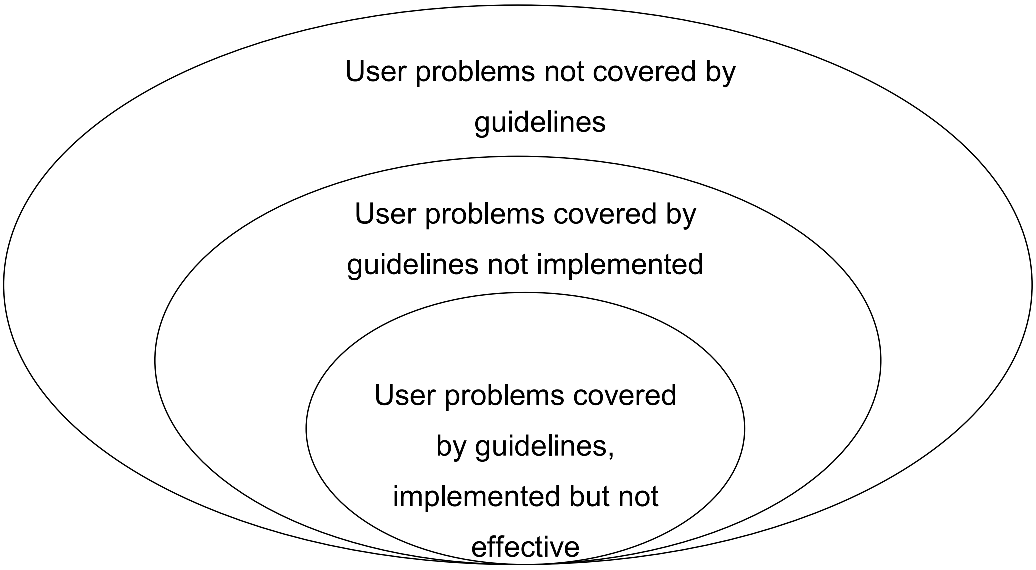


Figure 6.1– The overall set of user problems divided into three types: problems not covered by guidelines, those covered by guidelines but the guidelines are not implemented and those covered by guidelines with guideline implementations

As previously mentioned, the outer-most and the innermost sets are the ones that need special attention in analyses. It is very important to analyse the nature of problems that are not covered by guidelines, both for designers to know what they leave out when they evaluate the accessibility of websites, and for researchers to analyse the need for changes in current guidelines. The detailed discussion of the types of problems that were not covered by guidelines is presented in Section 6.2.5.

The innermost set of problems highlights very critical issues in relation to the effectiveness of guidelines in preventing user problems from happening. A detailed discussion of the types of problems not covered and the implications for design is presented in Section 6.2.4.

With regards to user problems covered by guidelines and not implemented in web pages, nothing can be said about the effectiveness of those guidelines. It could be that if developers had implemented those guidelines, user problems could have been prevented, but the results in this study cannot provide evidence to this. On another aspect, those problems reveal that many developers still do not implement existing accessibility guidelines in websites, possibly because of problems to understand the guidelines.

The amount of problems encountered in web pages that successfully implemented relevant checkpoint/success criteria is of great concern. It raises a very serious concern about the need to perform significantly more research to validate guidelines and techniques to implement accessibility features in websites, in order to develop them based on empirical evidence that supports their effectiveness. It also points that the use of un-validated techniques can be seriously misleading to developers, who may believe that the effort to implement a certain technique will improve the accessibility to disabled users, when they in fact will not necessarily do that.

As also stated in the report of the study performed by the Disability Rights Commission (2004), findings in this study reinforced the need to perform evaluation of websites with disabled users to uncover problems they can encounter on websites. Relying only on guidelines to evaluate web accessibility will most certainly leave many important accessibility problems uncovered.

6.2.4 Problems encountered by print-disabled users on websites not covered by WCAG 1.0 and WCAG 2.0

The results reported in Section 5.4 provided a detailed analysis of the types of problems encountered by print-disabled users that were not covered by WCAG 1.0 or WCAG 2.0.

WCAG 1.0 and WCAG 2.0 left a substantial amount of the problems encountered by print-disabled users uncovered. The cases where problems were not covered included both situations where there was no relevant WCAG 1.0 CPs or WCAG 2.0 SCs, or when one or more CPs/SCs had some relation to a given user problem, but it was not clear whether that it would effectively address the nature of the problem encountered by users in its entirety.

The first important observation about the large amount of problems not covered by guidelines is to highlight the importance of involving disabled users in the evaluation of the accessibility of websites. The results showed that, if only technical accessibility guidelines are used to evaluate websites, more than 50% of the problems can potentially be missed.

Some of the problems that were not covered by WCAG included issues that have been reported as usability problems listed in web usability guidelines (Nielsen 2000, Petrie and Power 2012, U.S. Department of Health & Human Services 2006). Problems not covered by web accessibility guidelines included issues with causes related to poor information architecture at website level ("navigation do not help users find what they are seeking" and "content not found where expected"), information architecture at page level ("irrelevant content before task content", "organisation of content is inconsistent with web conventions/common sense" and "too much information on pages"), system being too slow, and lack of feedback for actions.

It could be argued that these are not accessibility problems, but instead are usability problems and do not need to be addressed in WCAG. However, the research presented in this thesis came from the principle that web accessibility is about ensuring that people with disabilities can use the Web.

Previous research has shown that many problems are shared by disabled and mainstream users (Petrie and Kheir 2007). In that research, blind users reported significantly higher severity ratings than their mainstream peers for shared problems. Given the focus of the present study on disabled users only, it was not possible to perform analyses to confirm the results from this study.

General usability guidelines do not provide specific directions to help improve problems encountered by disabled that would be deemed as general usability problems. For example, besides considering a good design of dialogs in applications and providing good feedback messages, it is also important to think about how to ensure that messages are displayed in a way that can be used effectively by users with different disabilities and with different assistive technologies. If a good feedback message is placed where screen reader users cannot find them, those users will still struggle to have feedback from the system. Current guidelines for website navigation and organisation do not consider specific problems related to the cognitive load that blind users have when navigating in websites.

While specific web accessibility guidelines are not available to address those issues specifically to disabled users, it is important that designers be aware of their importance and take careful consideration of those problems when designing and evaluating websites to ensure they are accessible to disabled users. They need to be aware of the impact those issues may have on disabled users, and consider specificities that they may have that may involve different problems from other users, such as specific interaction methods and assistive technologies.

It is necessary to conduct more research into specific issues that the present study revealed were not covered by existing web accessibility guidelines. It is very important to explore into more detail the particular ways in which specific groups of users with disabilities may be affected by problems caused by issues such as information architecture, inadequate feedback, functionality problems, and others.

6.2.5 Problems encountered by print-disabled users on websites which successfully implemented guidelines

This study found very important and novel results showing that many problems encountered by print-disabled users can be found in pages that successfully implement the requirements to meet WCAG checkpoints/ success criteria. Those results were presented in Section 5.5.

The cases in which user problems occurred despite guidelines having been successfully implemented are very critical. Those cases highlight potential problems with guidelines providing ineffective solutions that do not address problems encountered by print-disabled users.

For WCAG 2.0, interface components were audited following the same process used for the selection of websites described in Section 3.2.3. According to this procedure, if a technique or set of techniques deemed as *sufficient* by the WCAG working group for a given success criterion was implemented, the success criteria was successfully implemented. However, in many cases where such “sufficient” techniques were implemented and users still encountered problems, this raised a serious question as to how “sufficient” the techniques really were.

Link destination to blind users

Being able to determine the purpose of a link is a crucial issue to blind users, as an essential aspect of navigating in web sites. It appears that relaxing the requirement to determine the purpose of a link in context as a level-A success criterion and leaving a more strict requirement for determining the purpose by the link text alone to a success criterion at level-AAA was not beneficial to blind users. This finding reinforces the question to the effectiveness of sufficient techniques. In many examples of problems in this study, the purpose of a link could be determined by the context of a link, such as a preceding heading, enclosing paragraph, or enclosing list, but users using a list of links out of context were not able to determine the purpose of such links.

The issue with links was further investigated in another study performed by Power et al. (2011), which also showed that not all techniques deemed as sufficient for SC 2.4.4 were equally effective in helping blind users.

Importance of audio description of video content

The different success criteria at different levels in WCAG 2.0 only made audio description mandatory if level AA of conformance is aimed. However, the option of providing a textual transcript as an alternative to achieve level A was rejected by users. This implies that, although having media alternatives other than audio descriptions of videos can be enough to achieve level A conformance, audio description is fundamental to provide blind and partially sighted users with the experience they expect when watching a video in a website.

As some users argued, having a textual alternative describing all scenes in a video can make them find general information in videos if they want, but do not provide them with the full experience they expect from a video. When content creators include videos in websites, they normally have a goal to transmit a message in a way that is better conveyed by a video. This goal would not be achieved if blind users are experiencing the video by reading a transcript only.

Providing alternatives to data in tables

Another interesting finding was that many blind users encountered problems when trying to read information from tables, even when tables had all the requirements recommended by guidelines.

The findings in this study showed that, even with screen readers that provide features to read tables and tables that conform to related guidelines, some blind users find it extremely difficult to manipulate tables, especially more complex ones. Even some users who rated themselves as somewhat experienced screen-reader users said they find it a very complex and demanding task to read information from tables.

Current guidelines seemed to be more effective for the most skilled screen-reader users, but failed to help other blind users that find the task of reading tables a difficult one itself. These results also showed that improvements need to be done in screen reader software to simplify their features to read tables, as many users considered it very difficult to use them.

Another interesting finding was that some partially sighted users (who did not read speech synthesis) also found it difficult to read tables. However, in their case, their assistive technology does not provide special features to help read tables with special mark-up recommended in accessibility guidelines as in the case of blind users. It can be very difficult for users with a high level of magnification to relate a table cell to its heading, when the screen can accommodate only a fraction of the table at one time. One recommendation from this finding is that more research needs to be done to elaborate accessibility guidelines that would make tables more accessible to partially sighted users, and to create features in assistive technology used by partially sighted users to help them read tables better.

Logical structure of headings for blind users

Some users encountered problems with headings structures that did not follow a strict nesting sequence, such as having a heading 2 before a heading 1. Making this stricter rule cease to be a requirement in WCAG 2.0 made many users encounter problems in pages that pass heading-related WCAG 2.0 success criteria.

For users who rely substantially on headings to navigate within web pages, having a heading structure that does not seem valid for them can seriously reduce their confidence on the headings on a page. They may be lead to believe, for example, that a page that starts with a heading at level 2 did not mark up the main heading, and that they might have lost something.

Clear identification of what controls or form elements do for blind users

In four problem instances, users could not identify the purpose of an input field even though it had an adjacent button that described its purpose, which is sufficient to pass WCAG 2.0 success criteria related to this issue. This happened mainly when users were reading a form sequentially, and they would only read the button if they continue reading the page after the input field. However, when arriving at an input field without a description, they became confused and did not continue reading the page immediately.

This indicates that providing labels associated with proper mark-up for input fields is a better solution for designers to use than relying on users exploring adjacent buttons.

Colour contrast for partially sighted users

Colour contrast is one of the key issues regarding the accessibility of websites to partially sighted users, and has been included in guidelines in both WCAG 1.0 and WCAG 2.0.

The results in this study showed that adhering to the requirements in different levels of WCAG 2.0 did not prevent partially sighted users from encountering problems in more than 70% of the problems that were covered by WCAG (which did not include images, as SCs 1.4.3 and SCs 1.4.6 only apply to texts and images of text).

The findings about colour contrast were very unexpected. This study design did not allow for more in-depth analyses to investigate and determine better contrast levels that would be more suitable to partially sighted users. However, this study did raise a very important gap pointing to the urgent need of considerable more research into what contrast levels are best for partially sighted users on websites.

Customising text colour for dyslexic users

As presented in Section 4.4, some dyslexic users may encounter problems to read text with black writing on white background. This can affect from very mild to very severe effects on how they read.

WCAG 2.0 SC 1.4.8 (level AAA) accepts that either an embedded colour selection tool is provided in a website or that the colours used can be changed by web browsers.

The results of this study showed that, despite having users who reported having problems to read black writing on white background in the panel, none of the users used a feature to change the background in their web browser. Those users reported a number of problems with websites that they expected would provide them with a built-in option to help them change colours. In eight problem instances, the web pages had passed SC 1.4.8 by implementing technique G156, as they allowed changing colours by the web browser. However, users still reported they wanted a different resource.

From another side, this also showed that there is a need to better train users with dyslexia to use available assistive technologies and adaptations available on operating systems and web browsers. However, in the case of web browsers, some users explicitly mentioned that they knew of the existence of features to change colours, but found it too complex to use them. This showed that there is need for more research to make it easier to use those adaptations in web browsers for disabled users.

6.2.6 The relationship between the severity ratings of problems encountered by print-disabled users and the priority levels of related guidelines in WCAG 1.0 and WCAG 2.0

This study yielded important findings showing the lack of correlation between the severity ratings of problems encountered by print-disabled users and the priority levels assigned to related guidelines in WCAG 1.0 and WCAG 2.0. The results from these analyses are presented in Section 5.6.

Priority levels of checkpoints and success criteria play a very important role to determine what accessibility requirements need to be implemented in websites to achieve certain levels of conformance to WCAG 1.0 or WCAG 2.0. For this reason, the lack of significant correlations between the priority levels of checkpoints related to problems encountered by users is very alarming.

For blind and dyslexic users, no significant correlations were found between the mean severity ratings of problems encountered by users and either the highest or the mean of the priority levels of related WCAG 1.0 checkpoints or WCAG 2.0 success criteria. For partially sighted users, significant correlations between severity ratings of user problems and priority levels of WCAG 1.0 checkpoints were found, but they were very low.

These results, based on a much larger set of data than other studies, confirm previous findings from studies performed by Petrie & Kheir (2007) and by Harrison & Petrie (2007) that pointed a lack of significant correlations between severity of problems encountered by users and the priority levels of WCAG 1.0 checkpoints. The results also showed that this problem still persists for WCAG 2.0, and that little improvement has been made with the new version of WCAG 2.0 with regards to priority levels better portraying the severity of problems to disabled users.

As described in Section 2.2, in WCAG 1.0, the statements that describe the priority levels seems to suggest that there would be some relationship between the priority levels and the difficulty disabled users would encounter if checkpoints were not addressed. For priority 1, for example, it is said that "[if a checkpoint is not satisfied], one or more groups will find it *impossible* to access information in the document", whilst for priority 2, it is said that "one or more groups will find it *difficult* to access information in the document", and for priority 3 that "one or more groups will find it *somewhat difficult* to access information in the document". In WCAG 2.0, levels of success criteria are determined according to a set of factors, which includes whether they are *essential*, or as described in WCAG 2.0. In other words, if the Success Criterion isn't met, then even assistive technology can't make content accessible" (Caldwell et al. 2008). Other factors considered include whether it is possible to satisfy a success criterion for "*all websites and types of content*", if it can be "*reasonably achieved by the content creators*", whether it could affect the look and feel of pages and whether there could be workarounds.

The results for WCAG 1.0 show that empirical evidence does not support the claims made in the statements about priority levels. For WCAG 2.0, although such claims about the impact on users have been removed, the results from this study are very important to inform developers about the very little relationship between the levels of success criteria and how they impact disabled users.

The most important lesson to be learnt from these results is that, in order to prioritise repairs of accessibility problems in websites according to the impact they have on disabled users, obtaining the severity ratings of problems by user evaluation is the most reliable measure source. Priority levels in WCAG seem to take other issues more into account than impact on print-disabled users, such as technical aspects related to the difficulty to implement, applicability to different types of technology and possible design limitations.

6.3 Limitations

Although it would be desirable to have as wide a range of disabled users as possible, this study focused on a more limited set of users to which evidence has shown find more accessibility in websites. The focus of the study was on users with print disabilities, namely blind, partially sighted and dyslexic users. According to results from the study performed by the Disability Rights Commission (2004), blind, partially sighted and dyslexic users covered the widest range of problems encountered by different groups of disabled users. For this reason, the present study did not include other groups such as deaf, hard-of-hearing and physically impaired users.

The lack of measures of time on task was a limitation of the present study. The think-aloud protocol was adopted due to the focus on uncovering the problems users had from their perspective. Users were encouraged to report and rate problems as soon as they were encountered. Due to this reason, it was not possible to perform precise calculations of the time taken on tasks, since users were allowed to pause momentarily to describe the problems they encountered.

Another limitation of the study was the impossibility to match users in different user groups in terms of computing experience and expertise with assistive technology. However, even though some consideration had to be taken with regards to users' computer experience for the analysis of some problems, the results obtained are very relevant in that they report very important issues that real users encounter when they use websites and what the nature of the problems they encounter is. Although computer experience and assistive technology expertise could not be matched between user groups, it was very important that this study did not include only very experienced disabled users. Studies that only consider users who have very advanced command of computers and assistive technologies can overlook many issues that less experienced users encounter every day when using websites and that need to be dealt with.

In the sample of dyslexic users, most users were in the mild-moderate range of the dyslexia severity spectrum. This means that accessibility problems that affect users with more severe levels of dyslexia may have not been revealed in this study.

Despite all efforts to perform an extensive search for conformant websites to the sample of this study, as described in Section 3.2, few websites had home pages at higher levels of conformance to WCAG 1.0 and WCAG 2.0 were encountered. However, the comparisons performed between websites at different conformance levels, even though not as numerous as expected, were a substantial contribution to the field, as other studies that compared user problems with WCAG using websites at different conformance levels could not be found in the literature.

This study revealed some unexpected results, which included a very high rate of problems with colour contrast encountered by partially sighted users that were conformant to WCAG 2.0. However, as this was not foreseen before the study started, it was not possible to perform more detailed analyses in order to establish better levels of colour contrast that would be best for users. Such a study would demand more well-defined experiments involving a range of users with different vision conditions, and a range of interface components with different colour contrast levels to test and determine their accessibility.

6.4 Summary of the chapter

The discussions in this chapter presented how the research conducted and reported in this thesis addressed the research questions proposed, with the characterisation of the main accessibility problems encountered by print-disabled users on websites, and the relationship between user-based measures of accessibility and measures of technical web accessibility based on WCAG 1.0 and WCAG 2.0.

The chapter presented discussions about the characterisation of problems encountered by blind, partially sighted and dyslexic users. The results from the analysed whether users could complete their tasks on websites, the difficulty to perform their tasks, and most importantly, the nature of the problems they encountered on websites, the frequency those problems were encountered and how severe they were. The results from the investigation of the problems resulted in a detailed description of the types of problems, and an in-depth analysis of the most critical types of problems encountered by each user group.

Following, the chapter presented the how the secondary research question was addressed, with descriptions of the findings related to each sub-question. The findings included the lack of significant differences between the number of user problems in websites that were conformant to WCAG and websites that were not, particularly to partially sighted and dyslexic users, the limited relationship between user problems and measures related to the number of violations of checkpoints/ success criteria in WCAG. The findings also showed that a large percentage of problems encountered by users were not covered by WCAG. Of greater concern was the fact that many problems that were covered by WCAG occurred in web pages that had successfully implemented checkpoints/ success criteria, which were still ineffective to avoid the user problems. Finally, the chapter also presented findings showing the lack of relationship between the severity ratings of user problems and the priority levels of related checkpoints/ success criteria.

The next chapter presents the conclusions of this thesis, the main contributions of this work and lines of investigation for future work.

Chapter 7. Conclusions

This chapter presents the main conclusions of the work presented in this thesis. It presents an overview of the research conducted and how it provided original contributions to knowledge in the field of Human-Computer Interaction in the area of web accessibility. It presents the main findings, implications and recommendations and directions for future work that needs to be developed in the area.

7.1 Overview of the research

The research presented in this thesis was motivated for the need for a better understanding of the problems encountered by print-disabled users when attempting to use websites. It is very important that websites are made accessible in order for disabled users to be able to use them effectively. However, there were still few studies that provided empirical evidence of accessibility problems by means of evaluation of varied samples of websites by disabled users. The last major study on the accessibility of websites involving disabled users was performed by the Disability Rights Commission of Great Britain (2004).

The main goal of the study reported in this thesis was to investigate the accessibility of websites by means of user-based measurements from evaluations with print-disabled users on websites. By performing those measurements, it was expected that this work would contribute with a significant improvement in the body of evidence of the main types of problems encountered by print-disabled users on websites, with a detailed description of their nature, as well as the frequency that they happen and how severely they impact print-disabled users.

Given the wide use of technical web accessibility guidelines reported in the literature, in particular those in the Web Content Accessibility Guidelines (WCAG) 1.0 and 2.0, a secondary goal of the study was to investigate the relationship between the problems encountered by print-disabled users and measurements related to WCAG.

The primary research question proposed in this research aimed to investigate *what are the main characteristics of accessibility problems encountered by print-disabled* *users when attempting to use websites*, and the secondary research question was to investigate *the relationship between user-based measures of accessibility of websites and measures of technical web accessibility based on the guidelines defined in WCAG 1.0 and 2.0.*

In order to address those research questions, the research performed consisted of one larger study with a complex design involving a number of measurements. The core of the study was the evaluation of a set of 16 websites by a panel including 32 blind, 19 partially sighted and 13 dyslexic users. The evaluations were performed using a concurrent think-aloud protocol, and users were asked to provide severity ratings to the problems they encountered. In order to address the secondary research question, the 16 websites sampled had a range of conformance levels to WCAG 1.0 and WCAG 2.0, carefully selected by accessibility audits of the home pages of hundreds of candidate websites, to ensure the best variability in the conformance as possible. This had the goal to analyse if there was an influence of conformance on the number of problems encountered by users.

Following the evaluation with print-disabled users, analyses were performed to verify if users problems were covered by guidelines, and if covered, if the web pages where they occurred implemented relevant WCAG checkpoints/success criteria.

The study yielded 3,012 user problems that were classified into categories according to the nature of the problems as perceived by users.

7.2 Findings and contributions

7.2.1 The characterisation of problems encountered by print-disabled users on websites

The main contribution of the work presented in this thesis was a substantial corpus of 3,012 problems encountered by blind, partially sighted and dyslexic users. The number of problems encountered places the corpus of user problems in this study as one of the largest encountered in the literature. The study also presents evidence of problems encountered with multimedia and interactive technology that is used much more frequently at present than at the time the latest large studies were performed (Coyne and Nielsen 2001, Disability Rights Commission 2004).

An important novel contribution of the present work was the identification of the most critical problems encountered by blind users based on the frequency and severity, which had not been reported in the previous large studies encountered in the literature.

For blind users, the most critical problems were those related to controls or form elements that did not work with the keyboard, lack of audio description of videos, unhelpful navigation structures, unclear descriptions of what controls or form elements do and lack of or insufficient feedback of their actions.

For partially sighted users, the most critical problems were related to problems with the presentation of controls, form elements, images and text (mainly related to colour and size), unhelpful navigation and poor information architecture making users not find what they expect in pages.

For dyslexic users, the most critical problems were caused by poor information architecture, unhelpful navigation, lack of expected functionality, such as search and auto-complete features in forms to help with spelling, difficulties to make sense of content due to language and incomplete information and problems with malfunctioning of features.

The study also found that blind and partially sighted users are the most affected in terms of not being able to complete their tasks on websites, finding it difficult to perform tasks.

7.2.2 Analysis of the relationship between problems encountered by print-disabled users and technical web accessibility

From the investigation of the secondary research question, the study provided important contributions to the understanding of the relationship between problems encountered by print-disabled users and technical web accessibility guidelines. The study confirmed previous findings (Disability Rights Commission 2004, Rømen and Svanæs 2008, Rømen and Svanæs 2011) that a large amount of user problems were not covered by WCAG 1.0 and WCAG 2.0.

Most problems that were not covered by the guidelines were related to poor information architecture, functional issues, lack of or inappropriate feedback and design of dialog in the interaction, lack of options to change specific features in the presentation of videos, lack of aids to help with spelling and specific issues with the presentation of text for dyslexic users.

A novel contribution of this work was the analysis of whether web pages where problems occurred had implemented relevant guidelines or not, that had not been performed in previous studies. The study revealed that many web pages had successfully implemented checkpoints or success criteria at some level of WCAG 1.0 or WCAG 2.0, meaning that they were ineffective to avoid the problems encountered by users. The main types of problems where guidelines failed to avoid problems for blind users were related to links that only made sense when placed in context (such as paragraph, or list item), implementation of alternatives to video other than audio description, illogical heading structure, lack of alternatives to tables and form fields that only made sense after an adjacent button was read. For partially sighted users, the levels of colour contrast required in WCAG 2.0 revealed to be not enough for users, and some users also wanted audio description and alternatives to tables. For dyslexic users, the most common issue was the lack of features to select colours directly on websites, when users did not know or could not change it in their web browsers.

Regarding conformance to WCAG, the study did not find significant differences between the numbers of problems encountered in non-conformant and conformant websites for most user groups. The only case where a significant difference was found was with the comparisons of problems encountered by blind users in websites that were conformant or not to WCAG 1.0.

Despite the efforts to find websites at all levels of conformance, it was not possible to find websites at higher levels of conformance to WCAG 1.0 and WCAG 2.0. For this reason, the comparisons were only performed between non-conformant and websites at any conformance level. However, despite this limitation, the analysis of the impact of achieving any level of conformance in reducing the number of user problems had not been reported in any study found in the literature.

Another important finding was a confirmation from previous studies (Petrie and Kheir 2007, Harrison and Petrie 2007) that did not find strong correlations between the severity ratings of user problems and the priority levels of related guidelines.

7.3 Implications and recommendations

The findings from the study performed in this thesis have very important implications for practice and research in web accessibility. The need for a move to evidence-based approaches to web accessibility involving disabled users is essential to make progress in developing techniques to make websites that disabled people can actually use.

Researchers on web accessibility should strive for conducting more research into problems encountered by disabled users and into solutions that have empirical evidence to support their validity and effectiveness. When new design recommendations are proposed, it is very important that those recommendations be accompanied by strong empirical evidence of how they help disabled users.

Future versions of web accessibility guidelines should consider including guidance to avoid accessibility problems encountered by users that are currently not covered by them. Even problems that could be regarded as “general usability” problems that also affect mainstream users, such as those related to information architecture and feedback to actions, may have particularities that affect disabled users in particular ways and would need special guidance.

Practitioners that work with the design and evaluation of websites should make sure that their activities are based on design recommendations that bear strong empirical evidence to support the claims that they will help disabled users. Using design recommendations that do not have strong empirical evidence may mean that substantial efforts from developers and evaluators are not translated into websites that can be used by disabled users.

Besides using evidence-based design recommendations, the findings in this thesis reinforce the importance of involving disabled users in the evaluation of websites as early as possible in the development of websites. Although accessibility audits performed by experts can help uncover some problems, a substantial amount of accessibility problems are only uncovered by accessibility evaluations of websites with disabled users.

When prioritising repairs of accessibility problems in websites, the impact of the problems on disabled users (measured from the severity ratings of those problems and the frequency that they happen) should be the main factor considered to make decisions. Other criteria, such as priority levels of technical guidelines, may not be as effective in improving websites to disabled users as considering first the problems that are most critical to them according to their severity and frequency.

Assistive technology developers and web browsers should consider many findings in this research that showed that disabled users had difficulties using particular features, such as readings tables with screen readers and magnification software, or changing colour schemes and font settings on web browsers.

7.4 Future Work

The findings resulting from the results presented in this thesis raised very important research questions that need to be addressed to improve the understanding of how disabled users interact with websites and which solutions work best to make websites more accessible to them. This study is one of many more studies that are necessary to build a strong body of evidence-based explanation of web accessibility from evaluations with disabled users

A first gap that needs to be explored is to further the investigation of the problems encountered by other groups of disabled users that were not targeted at the present study, such as users with hearing and physical impairments. It would be important to expand on previous studies and understand in more detail the nature of the problems they find, as well as the frequency and severity of those problems, following a similar methodology to that used in the present work.

Much more research is also needed regarding specific issues related to the user groups involved in this study. For the specific case of dyslexic users, substantially more research is needed to understand the problems they encounter into more detail. This is necessary to propose new design approaches and evaluate them carefully with users to describe in which ways they can be more effective. In order to do this, it is important to perform a number of studies including users with a wide range of difficulties that can be related to dyslexia.

For partially sighted users, the findings in this thesis point to a very important issue with currently used guidelines for colour contrast for websites. More in-depth research into colour contrast combinations that work well for partially sighted users needs to be performed. Such research should involve psychophysical aspects of colour vision and tests of different colour schemes with users with a range of colour sight and conditions to provide designers with better guidance into how to use colours in their designs of websites.

From the findings in this study that showed that many existing techniques for web accessibility are not effective to avoid user problems, future research should be developed to perform extensive studies into the effectiveness of techniques to avoid problems from happening to disabled users. One such study was performed (Power et al. 2011) to test different techniques to describe links to blind users. This study showed that the use of remote evaluation with disabled users can be an effective method to test the effectiveness of different techniques.

Besides investigating problems encountered by disabled users, it is very important that future research provides more evidence about the way disabled users use websites, which strategies they use, and how those strategies can be used to inform designers to build better interfaces. This is one of the goals of the i2Web project (Inclusive Future Internet Web Services)[21], in which one of the research lines involves investigating strategies used by disabled users and older adults to use websites, in particular Web 2.0 applications, such as those with user-generated content, social networks, and media convergence, including WebTV.

7.5 Summary

This chapter presented the main conclusions from this work and the contribution it presented to the field of web accessibility with the characterisation of problems encountered by blind, partially sighted and dyslexic users. The conclusions reinforced the importance of building a stronger body of evidence of problems encountered by disabled users on websites to build more effective solutions to make websites that disabled users can use. The chapter also presented implications of the findings to research and practice in web accessibility, and also future research directions that need to be further explored.

[21] i2Web Project – Available online at http://i2web.eu, last accessed on 23/09/2012

Appendix A. Informed consent form

Phd study on web accessibility

This study is part of my Phd research. It is investigating the accessibility of a range of web sites with the view to creating better measures of accessibility.

With your permission we will record the session (both audio and video), so that we can study the problems you encounter in detail later. Only myself, Helen Petrie, Chris Power and David Swallow, as the researchers working on web accessibility in the HCI Group at York will be allowed to view the recording.

You will be asked to do a number of tasks on a number of websites, while talking through the problems you have with the website. Each time you encounter a problem, I'd like you to rate it on a scale from 1 (cosmetic problem only), 2 (minor problem), 3 (major problem) to 4 (catastrophic problem, I can't proceed or I'd give up at this point). Because we want the researchers to also rate the severity of the problems, I don't want you to speak out the ratings you give, but show me with your fingers, and I'll note it down. That way the other researchers won't be biased if they watch the video and rating the problems themselves.

After each task, I will ask you a couple of questions, and when we have finished with each website, I will also ask you a couple of questions about that website.

Do you have any questions?

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 interview/focus group is over and you have been debriefed, you will be asked to initial the three statements in Section B, to indicate your agreement.

Section A

I, ________________________________, voluntarily give my consent to participate in this study on web accessibility. 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 completely confidential. Only Andre Freire, Helen Petrie, Chris Power and David Swallow will have access to the data collected.

________________________________ ____________________

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:

I was not forced to complete the interview/focus group. Your initials:

All my questions have been answered Your initials:

Appendix B. Problem rating form

Website: Participant:

Task:

Problem No	Rating	Location	Comments
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			

Appendix C. User information

User information questionnaire – adapted from Harrison (2008) .

Web accessibility testing participant information

Participant code	
Participant gender	Male / Female
Participant age	
Visual impairment	None Totally blind Partially sighted Nature of residual vision Since birth/age acquired: Braille reader? Yes/No Expertise with Braille:
Hearing impairment	None Profoundly deaf Sign language user Partially hearing Nature of residual hearing: Since birth/age acquired:
Physical impairment	None Nature of physical impairment Since birth/age acquired:

Dyslexia	None Nature of the dyslexia
Any other disability	Nature: Since birth/age acquired:
Assistive technologies used	(1) Type: Model: Version: How long used? Expertise? (2) Type: Model: Version: How long used? Expertise? (3) Type: Model: Version: How long used? Expertise?
Enhancements for the web	(1) Type: Explanation: (2) Type: Explanation: (3) Type: Explanation:

How many hours per week do you spend using websites?	a. never uses b. 1-5 hours c. 6-10 hours d. 11-20 hours e. more than 20
How long have you been using the Internet (including using www, email, gopher, ftp, etc.)? (please circle only one)	a) Less than 6 months b) 6-12 months c) 1-3 years d) 4-6 years e) 7 years or more
What is the main Internet browser you use? (please circle only one)	a) Internet Explorer b) Mozilla c) Firefox d) Opera e) Netscape f) Safari g) Don't know h) Other ____________________
What is your level of computer experience? (please circle only one)	None at all 1 / 2 / 3 / 4 / 5 / 6 / 7 Extensive
Have you ever participated in any website testing before? Yes / No (If yes – please state below how many times)	
What is your highest educational qualification?	a) Secondary / High School b) University c). Trade Qualification d). Other ______________________________

What is your native language?	a) English b) Other __________________
What is your employment status? (please circle only one)	a) Student b) Fulltime c). Part-time d). Self-employed e). Unemployed f). Home maker g). Retired
Would you be interested in taking part in other studies with the University of York?	a) Yes b) No
If yes, how would you prefer to be contacted to be informed about the studies?	☐ E-mail ☐ Telephone

Appendix D. Description of the main accessibility problems encountered by print-disabled users

This appendix presents a description of the main types of problems encountered by each user group. The description contains explanations of the nature of problems from the users' perspective and the main technical causes of those problems.

1. Description of the main problems encountered by blind users

This section presents a description of the main problems encountered by blind users, focusing on the problems that occurred most frequently and on the most severe problems. Firstly, the problems that were listed both as most frequent and most severe are presented. The section proceeds with the presentation of the problems listed as most severe only and the problems listed as most frequent but with low severity.

1.1. Problems encountered by blind users with high frequency and high median severity

This section presents the description of the 7 subcategories of problems encountered by blind users that were listed in the 15 most frequent subcategories and that median severity rating 3 (major) or 4 (catastrophe).

- ***Subcategory: Navigation elements do not help users find what they are seeking (navigation)***

Frequency: 99 instances (7.16%)

Median severity rating: 3 (major)

User problem

Users found that the navigation elements were confusing and disorienting, and did not help them find the information they were seeking in their task. In one example of problem in this sub-category, users were seeking the name of a cabinet minister in charge of public health in the Department of Health. The navigation several options that seemed to be plausible, such as “Public Health”, “About us”, “Contact”, but users could not be sure which one to follow.

Figure D.1 shows the navigation bar of the Department of Health website with the options available. In this example, the information about the referred minister was under About us/ Ministers.

Figure D.1 – Example of navigation of the Department of Health website – users had difficulties finding where to find the cabinet minister in charge of Public Health

Technical causes

In most cases when problems in this subcategory occurred, there were not specific problems with individual links available in the navigation. The reason for the problems was with the overall structure of the navigational elements, caused by a poorly designed information architecture.

- ***Subcategory: It is not clear what particular controls or form elements do (controls, forms or functionality)***

Frequency: 79 instances (5.79%)

Median severity rating: 3 (major)

User problem

Users encountered form elements or controls and could not determine what they would do. Examples include cases where users encountered buttons that read "unlabelled 1", or form fields that had labels that were not meaningful to users, such as "A-Z", or even form fields that gave users no label at all.

Technical causes

The main cause for problems in this subcategory was the use of unclear labels to identify controls and form elements, or the lack of labels or identification of those elements. The problems occurred with several types of elements, including HTML form elements, such as input fields, combo boxes, check boxes or buttons, and also with other interactive technologies such as Flash buttons.

In the case of HTML elements, many problems were related to the lack of a properly defined <*label*> element explaining the purpose of <*input*> elements, or <*label*> elements that did not explain the purpose of <*input*> elements properly. With components that used Flash technologies, many components had descriptors that were left with pre-defined values such as "unlabelled 1", "unlabelled 2", etc.

- ***Subcategory: No/insufficient feedback to inform that action has had an effect (controls, forms or functionality)***

Frequency: 72 instances (5.21%)

Median severity rating: 3 (major)

User problem

Users performed an action on the website and could not identify any feedback that the action had been performed. Problems included situations in which users activated a button or a link, and did not have any if the action had had any effect. In many of these cases, their screen reader remained silent after performing an action.

Other examples included cases where some message was given, but it was not sufficient for users to recognise that the action had been completed. For example, in a city council's website, users searched for local services based on their address given by house number, street name and postcode. In the next screen, users encountered the message "select address", followed by a list of addresses, in case there could be more than one address under the same number (in a block of flats, for example). When reading this message, users did not recognise this as an indication that their action of informing the address had been completed.

Technical causes

More than half of the problems in this subcategory was caused by the use of dynamic client-side features implemented on websites, such as features with Javascript or Flash. In one example, users activated a link named “Change location” in a ticket selling website, and the form to perform the action was included dynamically on the same page without reloading the current page on the browser. As this triggered no action on the browser, users did not know that anything had happened. Figure D.2 illustrates the screen with the new content added dynamically on the page after activating the “change location” link on this website.

Figure D.2 – Example of new content opened dynamically in the same page with no noticeable feedback to blind users

- ***Subcategory: Functionality does not work (as expected) (controls, forms or functionality)***

Frequency: 48 instances (3.47%)

Median severity rating: 3 (major)

User problem

Problems in this subcategory had to do with functionality that did not work or did not work in the way that users expected. Examples included cases where users expected a search feature to locate an item they were sure was on the website, or cases users expected a “sort” feature to show a list of items in order, but the sort feature showed users a list with fewer items than showed before using the feature.

Technical causes

Faulty implementation of functionality of websites were among the main technical causes of problems in this subcategory, such as in the example of a sort feature that lists a reduced number of elements after it is applied. Search features that did not find information users were sure should be on a website caused a substantial amount of problems listed in this category. The apparent cause of the problems in these cases was that some search features did not index all pages included in a website. In many such cases, websites only indexed pages included in a database of pages edited using a content management system, but did not include static pages.

- ***Subcategory: Control or form element cannot be reached using the keyboard (controls, forms and functionality)***

Frequency: 44 instances (3.18%)

Median severity rating: 4 (catastrophe)

User problem

Problems in this subcategory occurred when blind users were unable to have access to a control or form element using the keyboard. In many cases, for example, users expected that there should be a button somewhere when they detected that a form had ended or when they were aware of the existence of an interactive component on the screen, but were not able to get access to the element.

In one example, users tried to filter the search for a car in a vehicle manufacturer's website. In one example, users were looking for a video in a governmental website. Users went up and down in the page using the keyboard, going past the place that elements in the page seemed to suggest where the play button would be located, but they could not reach any button. In another example, users were trying to refine the search for a car in a vehicle manufacturer's website by budget. They found a text informing where they could select the "budget", but did not have access to the budget selector.

Technical causes

In most of the cases where this problem occurred, controls or form elements were not implemented accordingly to allow access via keyboard. Examples included Flash buttons that could only be activated using a mouse, such as in the cases of embedded videos that could not be played using the keyboard only or cases where controls were implemented using JavaScript that only allowed access using a mouse, such as in the example illustrated in Figure D.3.

Figure D.3 – Example of control – budget selector in a car manufacturer's website - that is not reachable using a keyboard

- ***Subcategory: No enhancement to audio, video or multimedia (audio, video or multimedia)***

Frequency: 31 instances (2.24%)

Median severity rating: 4 (catastrophe)

User problem

Blind users needed to find specific information in audio, video or multimedia, but were unable to get all the information they needed due to the lack of an enhancement, such as audio-description. In one example, users had to find information in a video in a museum website combining what was contained in audio and information that was only shown visually on the screen. Due to the lack of audio-description of graphical information, they were not able to obtain all information involved in their task.

Technical causes

The main technical cause for problems encountered by blind users in this sub-category was due to the lack of audio-description of visual information contained in videos.

- ***Subcategory: Expected functionality not present (controls, forms or functionality)***

Frequency: 31 instances (2.24%)

Median severity rating: 3 (major)

User problem

In many problems, user expected that a certain functionality would be present, but it was not. A very frequent example was in cases where users encountered a text box where they would enter text. The text box already contained some text (usually explanatory), and users expected that once they had started to type something in it, the box would clear itself up, but it did not. Another frequent problem was the lack of a search feature. Many blind users get accustomed to having a search feature on websites and often use the search as their first attempt at trying to locate information on websites. Not having a search feature available was reported as a problem for those users.

Technical causes

Not implementing an expected feature or functionality that was expected by users was the main cause of problems in this subcategory. The websites did not include features that were expected by users according to their mental model of websites.

1.2. Problems encountered by blind users with high median severity

- ***Subcategory: Audio content too difficult to understand due to background sound (audio, video or multimedia)***

Frequency: 5 instances (0.36%)

Median severity rating: 4 (catastrophe)

User problem

All problems in this category were also related to the task of booking a ticket for an event on TicketMaster. The website provide an alternative audio-captcha for users who had difficulties recognising characters in a distorted image. However, users found it very difficult to identify letters and numbers that were spoken very low and quickly in an audio with very loud and noisy background sound. This made it very difficult or impossible for users to recognise the letters and numbers to solve the captcha.

Technical problem

Audio captcha provided as an alternative to visual captcha had background noise that made was too difficult to recognise for users.

- *Subcategory: System times out (controls, forms or functionality)*

Frequency: 3 instances (0.22%)

Median severity rating: 4 (catastrophe)

User problem

All three instances of problems occurred in a task when users attempted to book tickets for an event. The last stage of the ticket booking process was solving a "captcha" to prove that the booking was done by a person and not by an automated system, by means of typing the text contained in a distorted image. Blind users used an alternative that consisted of recognising letters and numbers in an audio with a noisy background. Only three users managed to complete the "audio-captcha". However, the system had a time limit of 5 minutes, and users took much longer than this to solve the captcha. When they finally completed the task, they received a message that the system had timed out. Users would have to start the process all over from the start.

Technical causes

The technical cause of this problem was the limited time allowed for users to complete their booking process. The developers had not taken into account that disabled users could take considerably longer to complete the task.

- *Subcategory: No indication of how to interact with functionality (controls, forms or functionality)*

Frequency: 22 instances (1.59%)

Median severity rating: 3 (major)

User problem

Users get confused with how to interact with specific features due to interfaces not being clear about its operation or lack of instructions. Differently from the subcategory "it is not clear what particular controls or form elements do", problems in this subcategory had to do with users not being able to understand how the overall interaction with the functionality on a website works. Examples include a feature on the Digizen website where users could select two statements from a list to add to their profiles. The statements were organised in two columns, and the instructions were visually organised in two columns. However, users could not have any indication of where the columns were and could not understand how to interact with the functionality in order to accomplish their goal of creating their profile on the website.

Technical causes

In many problems encountered by blind users in this subcategory, the cause of the problem was related to the use of interface components that were not fully recognised by a screen reader. In the example mentioned in the Digizen website, the instructions mentioned elements that were visually in columns in a Flash interactive application. However, there was no indication other than the visual disposition that could be recognised by the screen reader to indicate which elements were in which column. In another example, the TicketMaster website did not use the <*select*> element to code a selection box. The functionality was emulated by Javascript that simulated the behaviour of a selection box, and visually, it looked like a selection box. However, screen readers could not recognise that it worked as a selection box and blind users did not know that they could operate it in the same fashion as they would operate a selection box.

- ***Subcategory: No alternative to text in specific format (text)***

Frequency: 17 instances (1.23%)

Median severity rating: 3 (major)

User problem

Users encountered problems with text in specific formats. With specific formats, users could not use the same strategies to navigate in the text as they did with regular HTML text.

Technical causes

All problems in this subcategory happened with users that had some aversion to text in PDF format. In some cases, the PDF documents had implemented accessibility features, but users still complained that they could not use the text in the PDF files linked from websites in the same fashion that they would use straight text on HTML pages.

- ***Subcategory: No/insufficient feedback that users have not performed a required action (controls, forms or functionality)***

Frequency: 16 instances (1.16%)

Median severity rating: 3 (major)

User problem

Users encountered problems in which they did not know that they had to perform a required action. This often occurred when the feedback message or instruction encountered by users was not clear to them or when they could not encounter a feedback message or instruction at all.

Technical causes

Many cases of problems in this subcategory involved unclear feedback messages or instructions. For example, in a city council's website, users searched for a local service by inputting a post ode and house number. The following page showed a list of possible addresses preceded by a message "Select address". In the cases reported, there was only one address in the list, and the message given was not clear enough about the action that users had to perform.

In other cases, feedback messages, often indicating incomplete required fields, were placed dynamically next to the field that needed to completed without reloading the page. No indication was given to users to inform that new content had been included in the page, as their screen readers remained silent.

- ***Subcategory: No alternative to audio, video or multimedia (audio, video or multimedia)***

Frequency: 15 instances (1.08%)

Median severity rating: 3 (major)

User problem

Users could not have access to multimedia content that was not accessible to them. In many of those cases, users' screen readers only read "Flash movie start" followed by "Flash movie end". Users became very frustrated to find that there was content on the screen that was not accessible at all to them.

Technical causes

The main cause of this problem was the lack of accessible alternatives to multimedia content or accessibility features, especially when using technology such as Flash. When not properly marked up, the use of such technologies may result in content being completely inaccessible to screen-reader users.

- ***Subcategory: Users cannot understand sequence of interaction (controls, forms or functionality)***

Frequency: 14 instances (1.01%)

Median severity rating: 3 (major)

User problem

Users encountered problems when a task involved a sequence of actions or steps had to be performed in a certain order, but the order was not clear to them. In one example, users tried to compare different cars available on a website and tried to use the "compare" button. However, they could not understand that they had to select up to three cars to compare before asking to compare them.

Technical causes

The interaction model of interfaces where those problems occurred were not logical to users and did not have clear indications of the step-by-step process that should be taken to complete tasks.

- ***Subcategory: Users inferred the existence of functionality where there was not one (controls, forms or functionality)***

Frequency: 12 instances (0.87%)

Median severity rating: 3 (major)

User problem

Users encountered parts of the page that looked like they had some functionality, mainly statements in the imperative form, such as "Book a redelivery". This made users believe that the specific piece of text would work as a link or a button, but when they tried to press Enter on the text, it did not act as an activator of any functionality.

Technical cause

Text with statements in the imperative form that looked like activators of some functionality were not turned into a link or a button, as would be expected by users.

- ***Subcategory: Inconsistent navigation structure in different pages (navigation)***

Frequency: 10 instances (0.72%)

Median severity rating: 3 (major)

User problem

When navigating in inner pages of a website, users encountered navigation structures that were inconsistent with the navigation in other pages of the website. Once users had already developed a mental model of the navigation structure available on the website, the inconsistency left them confused and disoriented.

Technical cause

The main cause of this problem was the lack of consistency in the navigation structure. In many websites, features such as search engines or special sections of websites are designed with a completely different navigation structure from the rest of the website.

- ***Subcategory: Broken link (underlying system characteristics)***

Frequency: 10 instances (0.8%)

Median severity rating: 3 (major)

User problem

Users encountered links that led them to an error page.

Technical cause

The URL (Uniform Resource Locator) of the destination in a link was wrong or destination page did not exist any longer.

- ***Subcategory: Language too complicated for perceived target audience (content)***

Frequency: 9 instances (0.65%)

Median severity rating: 3 (major)

User problem

Users found it difficult to understand the language level in pages related to their tasks, including difficult terms or difficult grammatical structures. In a legal website targeted at disabled people, users were seeking an explanation of the term “structured negotiation”. One would expect that the FAQ (Frequently Asked Questions) would not be targeted at experts, but users encountered an explanation with terms they found difficult, such as “advocacy” or “litigation”.

Technical causes

Use of difficult language without proper alternatives or explanations in a difficulty level that could be understood by non-experts in the topic of the website.

- ***Subcategory: Multimedia starting automatically is irritating (audio, video or multimedia)***

Frequency: 9 instances (0.65%)

Median severity rating: 3 (major)

User problem

Users got very irritated when audio content started to be played automatically when a page was opened. The sound played by the multimedia in the page made it very hard for users to listen to their screen-reader and interact with the page.

Technical causes

Embedded multimedia content on a page started automatically without the user knowing, jeopardising the users’ interaction with the page and competing with the screen reader.

- ***Subcategory: Link destination not present (link)***

Frequency: 8 instances (0.58%)

Median severity rating: 3 (major)

User problem

Users encountered links that did not inform their destination. In many such cases, users only heard “link – graphic”, and could not determine what the destination of the link was.

Technical causes

The main cause of problems in this subcategory was the use of images to identify the destination of a link and the complete absence of any textual alternative to the images with the *alt* attribute in the <*img*> tag.

- ***Subcategory: Default presentation of text not adequate (text)***

Frequency: 6 instances (0.43%)

Median severity rating: 3 (major)

User problem

Users found it difficult to make their screen readers read text that was not in an appropriate presentation format. The majority of the problems encountered by blind users related to this issue was on an educational website with information about cyberbullying. Users needed to find information about the authors of a research study related to statistical information shown in the text, but they had to go through too much effort to have the screen reader read it. The text was only shown graphically when the text was hovered over with a mouse or when a special setting on the screen reader was activated. Figure D.4 illustrates the text that shown as a “tooltip” associated to the information “22% of children and young people...”.

Figure D.4 – Example inadequate presentation text as a “tooltip”

Technical causes

The text users were seeking was coded using the *title* attribute, associated to an excerpt of the text as in "*22% of children and young people...*. Screen-readers do not read this automatically, unless users adjust a special setting.

- ***Subcategory: Navigation elements not understandable (navigation)***

Frequency: 6 instances (0.43%)

Median severity rating: 3 (major)

User problem

Users could not understand the meaning of jargon and difficult words in the navigation of websites. When looking for a painting in a museum's shop website, users encountered words such as "Pre-Raphaelite" and "antiquarian", for example, which made it difficult for them to choose a link and continue with their task.

Technical cause

The cause of those problems was the use of jargon and difficult words as navigation elements.

- ***Subcategory: Users cannot associate table cell with headers (tables)***

Frequency: 5 instances (0.36%)

Median severity rating: 3 (major)

User problem

Users encountered problems with long tables where they needed to relate the information in a table cell with a table header. Blind users were not able to establish the meaning of the content in a table cell without the content of the header when the relationship between the cell and the header was only visual.

Technical causes

No relationship between a table cell and its corresponding heading was made in the source code. In HTML, the <*td*> element did not use the *scope* or *header* attributes to relate a cell in a table to its corresponding headers.

- ***Subcategory: No alternative to functionality (controls, forms or functionality)***

Frequency: 3 instances (0.22%)

Median severity rating: 3 (major)

User problem

Users encountered functionality that was only available in technology that they find difficulties using, such as Flash, and no alternative with a different technology is provided.

Technical causes

Developers only made available a version of functionality with technology that users find difficult to use.

- ***Subcategory: Inability to change presentation of text (text)***

Frequency: 3 instances (0.22%)

Median severity rating: 3 (major)

User problem

Blind users encountered problems with text that they could not understand properly when spoken by the screen reader, and could not change the way they were spoken by any means. Users could not understand words when they were spelled letter by letter, such as I – N – T – O – U – C – H, for “in touch”.

Technical problems

Words with extra spacing used for formatting caused problems for screen-reader users.

- ***Subcategory: Inadequate alternative to functionality (controls, forms and functionality)***

Frequency: 3 instances (0.22%)

Median severity rating: 3 (major)

User problem

Users were having difficulties to use a service to find local services in a city council's website due to poor feedback and instructions. In an attempt to use a different feature, users tried the "accessible version" option available on the website. However, they were disappointed to find that the accessible version offered little changes in relation to the original version. The original version of the search had three form fields: - house number, street name and postcode, whilst the accessible version had only one form field called "address", where users could enter their address all in one field. Users complained that they had no problems with the form fields with the original version, and that the only change they could see offered them no improvement in relation to the problems they were having.

Technical causes

A version named "accessible version" of a form field to search local services offered no improvements in relation to the original version.

- ***Subcategory: System executes action unexpectedly (controls, forms or functionality)***

Frequency: 3 instances (0.22%)

Median severity rating: 3 (major)

User problem

Problems occurred when users tried to see use a list of items in a selection box. They expected that they would be able to use the down arrow and see the list of items available, decide which item to choose and then perform an action. However, as they pressed a button to see the next item, the system automatically selected the item and performed the action to continue unexpectedly.

Technical causes

The selection boxes used a Javascript action associated to pressing a key on the keyboard. This made it impossible for blind users to list all items in the box without performing an action, since they needed to use the keyboard to go to the next item.

1.3. Problems encountered by blind users with high frequency and lower median severity

- ***Subcategory: Link destination not clear (link)***

Frequency: 117 instances (8.46%)

Median severity rating: 2 (minor)

User problem

Users encounter links with a destination description that is unclear to them, be it a non-sense text or text that they can make sense, but does not indicate clearly the destination of a link. This could happen either when users encountered a link in a sequence when reading the text of a page, or when users used a feature from their screen reader to list all links available on a web page.

Technical causes

The use of images with inappropriate alternative texts to indicate the destination of a link was one of the most frequent causes of problems with unclear destinations of links for blind users. Many links contained destinations that had non-sense text, including very long codes that did not make sense or file names. Those were images that contained the description of the destination visually, but failed to provide an adequate textual alternative. Problems caused by links with images without textual alternatives accounted for 30.8% of the 117 problems in this subcategory. Other causes include link destinations that did not make sense out of context, such as “read more” at the end of a paragraph or numbers marked as links in a list of search results, which accounted for 22.2% of all problems in this subcategory.

- ***Subcategory: Content not found in pages where expected by users (content)***

Frequency: 88 instances (6.36%)

Median severity rating: 2 (minor)

User problem

This subcategory describes problems where users confidently followed a link to a page, but a piece of information that they expected to find there was missing. For example, on a governmental website, users sought information about the name of a cabinet minister and followed the link "Contact us". They expected that the page would list all names and contact information of the ministers, but they could not find any information about the specific minister they were seeking.

Technical causes

The main technical cause of problems in this subcategory was connected to the information architecture and organisation of websites. Users expected to find certain pieces of information in parts of the website that were different from that laid-out by the designers of the websites.

- ***Subcategory: Irrelevant content before task content (content)***

Frequency: 87 instances (6.29%)

Median severity rating: 2 (minor)

User problem

Users often encountered problems with large blocks of content irrelevant to their task occurring before the relevant content. For example, when users were seeking information about insurance plans, the relevant page had lengthy descriptions of why it was important to buy insurance before a summary of insurance plans, the relevant content on the page.

Technical causes

The order of content in the page did not follow the logical order that would be expected by users, with the most relevant content appearing first. Instead, developers placed advertisement content first, making users spend longer to get to the relevant content they needed.

- ***Subcategory: Users cannot make sense of content (content)***

Frequency: 66 instances (4.77%)

Median severity rating: 2 (minor)

User problem

Users encountered problems with content they could not make sense. Examples include text that was placed out of context, such as "Silver Torc L" as the only content in a web page about an exhibit in a museum. Other examples include cases where users could not make sense of acronyms, abbreviations and specific terms they were not familiar with.

Technical causes

The lack of explanations for acronyms and abbreviations was the cause of a substantial number of problems in this subcategory. The use of words out of context was also the cause of many problems. In many such cases, the text that was difficult to make sense was an alternative text to an image that was not placed in the context of the rest of the page, as in the example of the "Silver torc L" in the museum website. Many problems were related.

- ***Subcategory: Destination not what was anticipated (navigation)***

Frequency: 48 instances (3.47%)

Median severity rating: 2 (minor)

User problem

Users encountered problems when the destination of a link was not what they expected. In several examples, users activated the link 'Home' on a web page expecting to go to the home page of the website. However, they arrived at the a page describing a subsection of the website.

In other examples, users expected that a link would be opened in the same window of their web browser, but it opened it a new window. For many users, having a new window opened caused some confusion, as not all users remembered that they were not at the same window that they were when they started to navigate in the website.

Users also encountered problems when using links to a part of the same web page where they were. For example, users tried to use a link "skip to content" to make their screen readers skip the navigation bar and go straight to the main content in the page. However, they still remained at the same place even after using the link.

Technical problems

Different issues caused problems in this subcategory in different occasions. A common cause of problems was the separation of content into different related websites. For example, the website “NHS Direct” had links to health information that was contained in the website “NHS Choices”. After navigating in the NHS Choices website coming from the NHS Direct website, users often expected that the “Home” link would take them to the original home page of the website where they started.

Another common cause of problems was not informing users that a link would open in a new window. Blind users need to know when a new window is open, since this impacts on their awareness of how many and which windows are open.

The problem with internal links to skip straight to the main content was associated with poorly coded link destinations and mark-up.

- ***Subcategory: No headings (headings)***

Frequency: 41 instances (2.96%)

Median severity rating: 2 (minor)

User problem

A substantial amount of blind users tried to have an overview of a web page by reading its heading structure. Users encountered problems when web pages had no heading elements at all. This made users have to use a different strategy to navigate from that they are most accustomed to, which implied in more effort and time.

Technical causes

Developers did not structure content in web pages divided by topics, or did not use proper mark-up in HTML to inform which elements should be interpreted as headings.

- ***Subcategory: Organisation of content is inconsistent with web conventions/common sense (content)***

Frequency: 39 instances (2.82%)

Median severity rating: 2 (minor)

User problem

Users encountered problems when content was not organised according to common sense or web conventions. In one example in a web designer's website, users were looking for comments from a client of a specific project. The projects page had a list of projects and a separate list of quotes from clients in a different client, and users found it difficult to relate the quotes to the project they wanted to know about. Navigating linearly with a screen reader made it even more difficult for users to find related content elsewhere on the same web page. Figure D.5 illustrates the example from the Green Beast website, with projects and comments about projects in non-related areas.

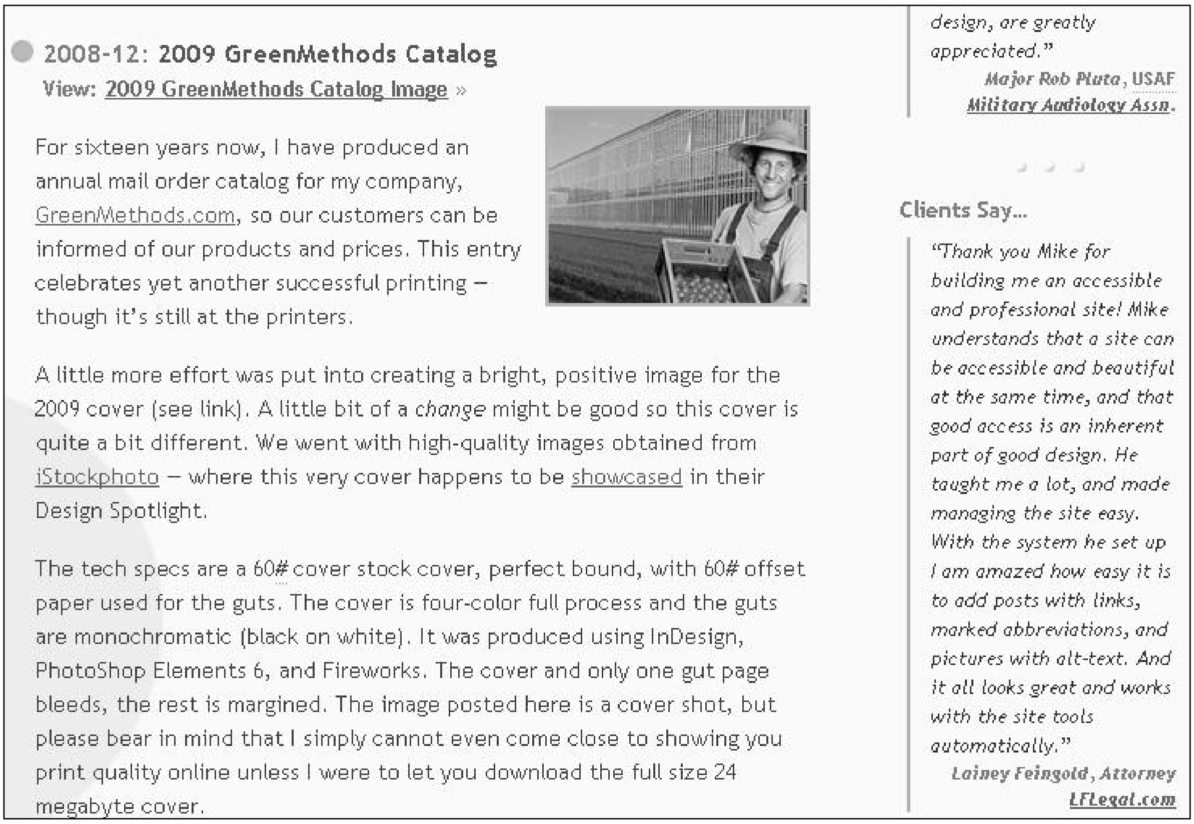


Figure D.5 – Example of website with bad organisation – comments about projects not located next to related project

Users also encountered problems with other issues related to the organisation of content, such as finding it difficult to search for a specific item due to a list not being in alphabetical order.

Technical causes

Poorly laid-out pages and lack of structure in pages were the main causes of problems in this subcategory. Content was not placed logically on a web page to make it easier for users to find.

- ***Subcategory: Inadequate alternative to image (image)***

Frequency: 33 instances (2.39%)

Median severity rating: 2 (minor)

User problem

Users encountered problems with images that had inadequate alternative text to describe its content. Examples included images that only read to users “graphic 1”, “graphic 2”, or some undecipherable text that users could not understand at all.

Technical causes

The use of inadequate text to describe an image was the main cause of problems in this subcategory. Developers included codes, file names or other names in the *alt* attribute of *<img>* tags.

2. Description of the main problems encountered by partially sighted users

This section presents a description of the main problems encountered by partially sighted users, focusing on the problems that occurred most frequently and on the most severe problems. Firstly, the problems that were listed both as most frequent and most severe are presented. The section proceeds with the presentation of the problems listed as most severe only. For partially sighted users, all problems listed as most frequent were also listed as major or catastrophic problems.

2.1. Problems encountered by partially sighted users with high frequency and high median severity

This section presents the description of 15 most frequent subcategories, which also happened to have median severity rating 3 (major) or 4 (catastrophe).

- ***Subcategory: Default presentation of text not adequate (text)***

Frequency: 157 instances (16.77%)

Median severity rating: 3 (major)

User problem

Users encountered a substantial number of problems with text that they had difficulties to read due to inadequate presentation. In problems in this subcategory, users would somehow be able to change the presentation of text by changing colour contras or the size. However, changing the settings would still result in a non-satisfactory presentation, or the necessary changes would demand too much effort from users.

Technical causes

The most frequent cause of problems with presentation of text, accounting for nearly 46% of problems in this subcategory, was bad colour contrast between text and its background. The colour used by web designers did not have enough contrast with the background in the default presentation, or became unreadable when special colour settings were applied by users.

Inadequate font size was another frequent cause of problems, accounting for approximately 18% of problems in this subcategory. Users complained that the font size was still too small even after magnification.

Other causes of problems also include use of text in images with low resolution that became blurred when magnified, or presentation of text in more than one column, that made users spend more time panning with their screen magnifiers.

- ***Subcategory: Navigation elements do not help users find what they are seeking (navigation)***

Frequency: 78 instances (8.33%)

Median severity rating: 3 (major)

User problem

Like blind users, partially sighted users also found that the navigation elements were confusing and disorienting, and did not help them find the information they were seeking in their task.

Technical causes

In most cases when problems in this subcategory occurred, there were not specific problems with individual links available in the navigation. The reason for the problems was with the overall structure of the navigational elements, caused by a poorly designed information architecture.

- ***Subcategory: Content not found in pages where expected (content)***

Frequency: 77 instances (8.23%)

Median severity rating: 3 (major)

User problem

Like blind users, partially sighted users also encountered problems when they confidently followed a link to a page, but a piece of information that they expected to find there was missing.

Technical causes

The main technical cause of problems in this subcategory was connected to the information architecture and organisation of websites. Users expected to find certain pieces of information in parts of the website that were different from that laid-out by the designers of the websites.

- ***Subcategory: Difficult to scan page for specific item (all media types)***

Frequency: 44 instances (4.7%)

Median severity rating: 3 (major)

User problem

Users found it difficult to scan for specific items in a page. In one example, users were trying to find the search box in a museum shop website, but they could not spot where the box was, as it was tangled in the middle of a lot of other content. In another example, users needed a link to a specific page that was in the middle of a paragraph, but they found it difficult to spot the link, as it did not stand out from the rest of the content.

Technical causes

The main cause of problems in this subcategory was the lack of effective means to help scan for items by means of structural elements and highlighting mechanisms. Problems occurred with items such as text, images, controls and form elements that did not stand out visually on the screen.

- ***Subcategory: Default presentation of control or form element not adequate (controls, forms or functionality)***

Frequency: 43 instances (4.59%)

Median severity rating: 4 (catastrophe)

User problem

Partially sighted users found it difficult or impossible to see or to interact with text or icons in controls or form elements due to inadequate presentation. In most cases, partially sighted users had to use different colour and size settings with assistive technologies or with settings in their operating systems. In one example, users had to select a seat to buy a ticket for an event by clicking on a circle on the seat map. The circle was very small, and some users complained that they "blended together". Most users also had problems to click on the circle due to its small size, which required a lot of precision from them. Figure D.6 presents a screenshot of this seat map from the TicketMaster website.

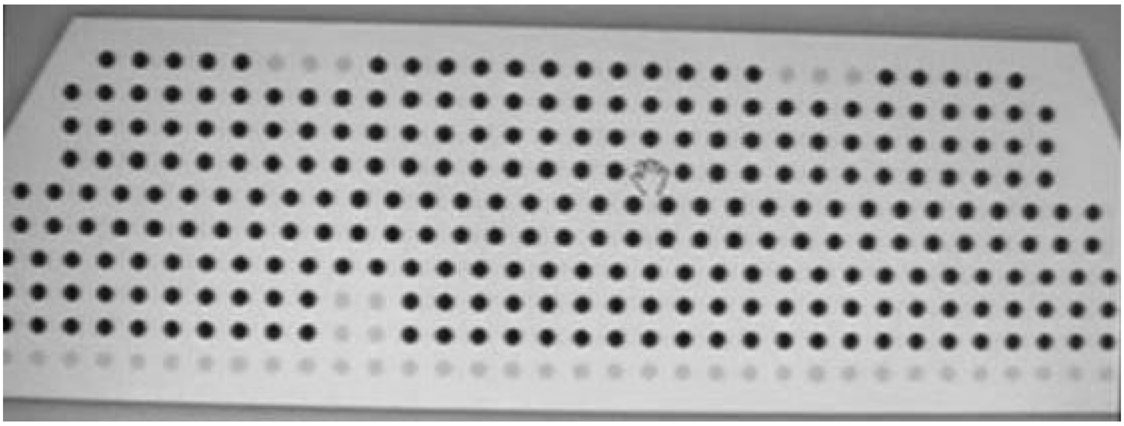


Figure D.6 – Example of control with inadequate presentation – circles to select seat in a seat map have low colour contrast and are too small

Technical causes

Problems in this subcategory were caused by poor colour contrast or small sizes of text and images in controls or form elements. Problems with colour contrast and size made it difficult for users to perceive where controls or form elements were or to identify their identification. Problems with small size and area of interaction also made it difficult for users to interact, since they required too much precision for users.

- ***Subcategory: Irrelevant content before task content (content)***

Frequency: 40 instances (4.27%)

Median severity rating: 3 (major)

User problem

Users encountered problems when irrelevant content was placed before task content. In the particular case of users with screen magnification software, the search for relevant content often left users lost and made them spend much more effort due to the panning required to explore parts of the screen. When magnification software is used, only a small part of the screen is shown at one time.

Technical causes

The order of content in the page did not follow the logical order that would be expected by users, with the most relevant content appearing first. Instead, developers placed advertisement content first, making users spend longer to arrive at the relevant content they needed.

- ***Subcategory: Expected functionality not present (controls, forms or functionality)***

Frequency: 33 instances (3.53%)

Median severity rating: 3 (major)

User problem

Partially sighted users also reported problems when they expected that a certain functionality would be present, but it was not. The most frequent expected feature by users was a search box. In a city council's website, users tried to use an accessible version of a search for local services after using the regular search. They expected that the system would have kept the data they had typed previously for the accessible version, but they had to type all information again.

Technical causes

Not implementing an expected feature or functionality that was expected by users was the main cause of problems in this subcategory. The websites did not include features that were expected by users according to their mental model of websites.

- ***Subcategory: Too much information on page (content)***

Frequency: 33 instances (3.53%)

Median severity rating: 3 (major)

User problem

Users became disoriented and overwhelmed when they pages had too much information on them. This cause special trouble to users with screen magnifiers, who had to spend considerably more time exploring different parts of the screen since they only had a small viewport.

Technical cause

The cause of problems in this subcategory was the excessive amount of information that was presented all at once in one web page. Those pages would often lack structural elements to break their parts into smaller blocks that would be easier for users to read.

- ***Subcategory: Default presentation of image not adequate (image)***

Frequency: 32 instances (3.42%)

Median severity rating: 4 (catastrophe)

User problem

Users encountered problems to see information in images when using screen magnifiers or when changing their settings to increase the size and change colour schemes. Users complained that images became blurred or pixelated when magnified. They also had problems when changes in colour background made it impossible to see what was in the image.

Another common issue encountered by users was the use of glary images on pages. Users found that images with bright white backgrounds gave them glare, often making them turn their faces due to the pain it caused or jeopardising their sight momentarily and preventing them from seeing what was shown around the glary image. In many cases, the image with glary white background would take a substantial part of the screen due to the magnification, which would increase even more the impact on users, who often had to use the computer when ambient lights switched off. Figure D.7 presents an example of an image from the Pret A Manger website with a bright white background.

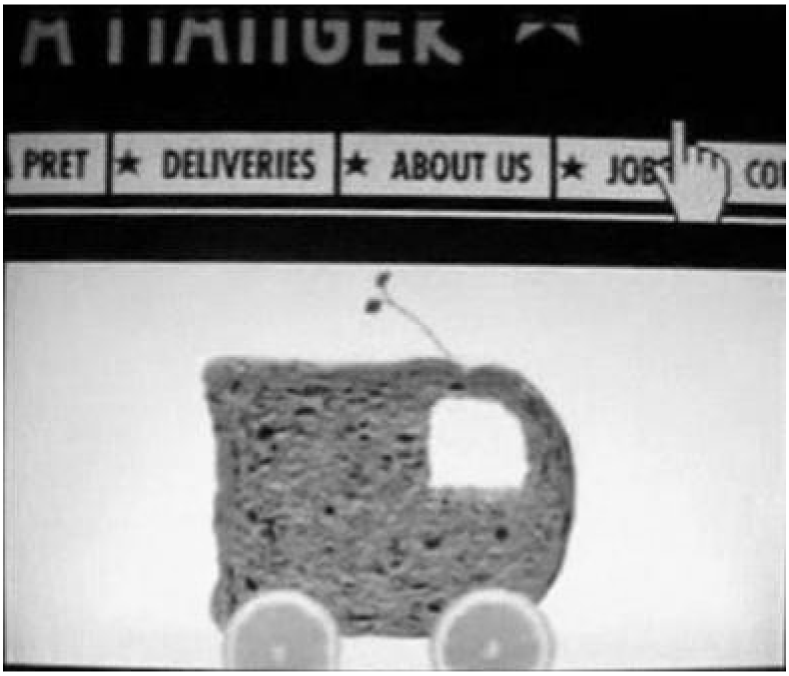


Figure D.7 – Example of image with bright white background that causes glare to partially sighted users

Technical causes

Problems with glary images occurred mostly with users that had to change their colour scheme. When the change was made using the web browser's settings, changes in colour depended on the layout specifications in the web page. If an image has an opaque white background, its background will not change even if users choose to have a black background, as shown in Figure D.7.

Images with low resolution are also the cause of problems with low readability of information in images. If the resolution is too low, images get pixelated when amplified.

- ***Subcategory: Inability to change presentation of audio, video or multimedia (audio, video or multimedia)***

Frequency: 31 instances (3.31%)

Median severity rating: 3 (major)

User problem

Users found it difficult to listen to audio or watch a video or animation due to the inability to change aspects of its presentation. In several cases, users complained because a video was too small for them to watch and they were not able to make the video larger or full screen. In other cases, users complained because subtitles or other text on videos or animations disappeared too fast from the screen and they could not change the speed in order to read it.

Technical cause

The main cause of problems in this subcategory was the lack of options to change the way in which audio, video or multimedia is presented. Particularly with videos and animations, there were not options available to change aspects such as size and speed presentation.

- ***Subcategory: Functionality does not work (as expected) (controls, forms or functionality)***

Frequency: 30 instances (3.21%)

Median severity rating: 3 (major)

User problem

Like blind users, partially sighted users also encountered problems with functionality that did not work or did not work in the way that users expected. Examples included cases where users encountered a search for local services at a page where schools were listed. They expected that the search would only return schools, but it returned a list of all sorts of services and local governmental information.

Technical causes

Faulty implementation of functionality of websites was among the main technical causes of problems in this subcategory. In the example on the city council's website, the cause of the problem was the loss of context of the search – users arrived at a search functionality from a list of schools, but the functionality did not take this into account when providing the list of services from the search.

- ***Subcategory: Organisation of content is inconsistent with web conventions/common sense (content)***

Frequency: 29 instances (3.1%)

Median severity rating: 3 (major)

User problem

Users encountered problems when content was not organised according to common sense or web conventions. In one example in city council's website, users were looking for information about local primary schools. In the primary schools web page, users were told to use the "find my nearest service". However, the users could not find the referred link easily where they would expect.

Technical causes

Poorly laid-out pages and lack of structure in pages were the main causes of problems in this subcategory. Content was not placed logically on a web page to make it easier for users to find. In the city council's example, the link "find my nearest" was placed in a separate column from the place where a reference to it was made. Figure

D.8 shows an excerpt of the screen with a list of schools in the leftmost column and a reference to the “find my nearest” link; the referred link was placed in the right-hand side column, which in some cases was out of partially sighted users’ viewport when they were using magnification.

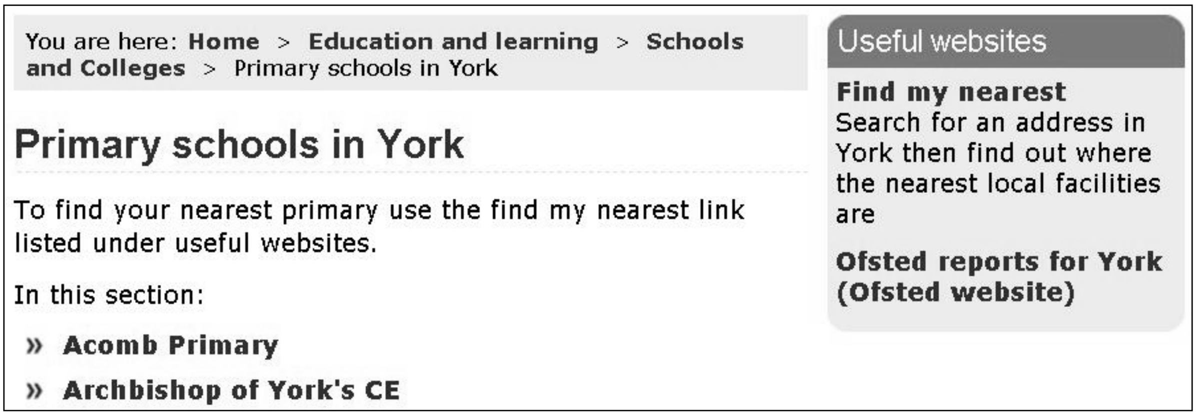


Figure D.8 – Example of illogical organisation of content – link “Find my nearest” not placed next to where it is referenced

- ***Subcategory: Users cannot make sense of content (content)***

Frequency: 28 instances (2.99%)

Median severity rating: 3 (major)

User problem

Users encountered problems with content they could not make sense. Most examples were related to users encountering information that was not put in context or did not have enough explanation to them. In one example, users encountered information about bus fares for park and ride services in a city council’s website. However, they were not sure if the fare applied to all bus lines or not, as there was not enough information to indicate this.

Technical causes

Incomplete information on pages was the main cause of problems in this subcategory for partially sighted users. In the city council’s web site, for example, the page had links to the bus lines with park and ride services available and the fares, but there was no explicit explanation that the fares applied to all lines.

- *Subcategory: Inability to change presentation of text (text)*

Frequency: 22 instances (2.35%)

Median severity rating: 3 (major)

User problem

Users encountered problems when they were not able to change certain aspects of the presentation of text with the available features on their assistive technologies or web browsers. Partially sighted users reported problems when they could not change the colour background from a glary white in a PDF document using the colour settings on Windows. Other examples include cases where users could not change the vertical orientation of text in a Flash interactive application, or change the spacing between two pieces of text in a restaurant's menu to be able to see product and price at the same viewport.

Technical cause

The main cause of problems in this subcategory was the use of technology that did not allow changes in the presentation of text or were not compatible with certain assistive technologies or operating system's settings. On PDF documents, for example, the colour background of documents did not follow the background settings set at the operating system. Flash applications did not allow changes in the presentation of text to be made by users.

- *Subcategory: Destination not what was anticipated (navigation)*

Frequency: 22 instances (2.35%)

Median severity rating: 2.5 (minor-major)

User problem

Like blind users, partially sighted users also encountered problems when the destination of a link was not what they expected. In several examples, users activated the link 'Home' on a web page expecting to go to the home page of the website. However, they arrived at the a page describing a subsection of the website.

In other examples, users expected that a link would be opened in the same window of their web browser, but it opened it a new window. In other cases, users got confused when the page where they arrived was not the page they expected by the name of the link where they clicked before. On the TicketMaster website, users were trying to buy a ticket for an event at the Grand Opera House in York. On a list of events, users encountered a link with the name of an event followed by a date in the text. They expected that the link would take them to a page to buy a ticket, but they arrived at a calendar page with a list of dates.

Technical causes

Different issues caused problems in this subcategory in different occasions. A common cause of problems was the separation of content into different related websites. For example, the website “The AA” had links to educational information that was contained in the website “AAttitude”. After navigating in the “AAttitude” website coming from the main “The AA” website, users often expected that the “Home” link would take them to the original home page of the website where they started.

Another common cause of problems was not informing users that a link would open in a new window. Like blind users, partially sighted users also need to know when a new window is open, since this impacts on their awareness of how many and which windows are open when they have a small viewport.

Many problems also were related to links having unclear destinations, especially when the context surrounding the links changed the interpretation that users had. In the TicketMaster example, although the link only had the name of the event, having the date next to it made the users believe that it would take them to a booking page. Figure D.9 shows the link in the TicketMaster website and its disposition, which suggested to users it would take them to a booking page for the event on that date and time.

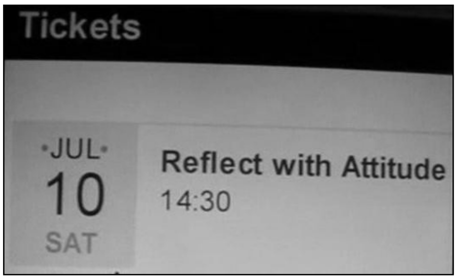


Figure D.9 – Example of link that did not take users to where they expect

2.2. Problems encountered by partially sighted users with high median severity

- ***Subcategory: Navigation elements not understandable (e.g. using jargon of difficult language) (navigation)***

Frequency: 6 instances (0.64%)

Median severity rating: 4 (catastrophe)

User problem

Like blind users, partially sighted users also find problems when they could not understand the meaning of jargon and difficult words in the navigation of websites.

Technical cause

The cause of those problems was the use of jargon and difficult words as navigation elements.

- ***Subcategory: Default presentation of audio, video or multimedia not adequate (audio, video or multimedia)***

Frequency: 5 instances (0.53%)

Median severity rating: 4 (catastrophe)

User problem

Users encountered problems when the default presentation of video and multimedia was inadequate, even in cases when they would be able to change them. In some examples, a video would be shown with a white background that could be changed using users' assistive technologies. However, showing the video with the white background unexpectedly caused a lot of glare to users, that would have to wait some time to have their usual sight restored.

Technical causes

Video content with white glary background and small window views were the main technical causes of problems in this subcategory for partially sighted users.

- ***Subcategory: Audio content too difficult to understand due to background sound (audio, video or multimedia)***

Frequency: 3 instances (0.32%)

Median severity rating: 4 (catastrophe)

User problem

Like blind users, some partially sighted users also encountered problems in the task of booking a ticket for an event on TicketMaster. The website provide an alternative audio-captcha for users who had difficulties recognising characters in a distorted image. However, users found it very difficult to identify letters and numbers that were spoken very low and quickly in an audio with very loud and noisy background sound. This made it very difficult or impossible for users to recognise the letters and numbers to solve the captcha.

Technical problem

Audio captcha provided as an alternative to visual captcha had background noise that made was too difficult to recognise for users.

- ***Subcategory: No enhancement to audio, video or multimedia (audio, video or multimedia)***

Frequency: 6 instances (0.64%)

Median severity rating: 3.5 (major-catastrophe)

User problem

Some partially sighted users with more severe sight loss encountered problems when no audio description of visual information in videos was provided. Users with little sight reported problems with lack of audio description in cases when they had to read text on videos or detailed parts of images.

Technical causes

Lack of audio description providing description of scenes and visual information contained in videos.

- ***Subcategory: No indication of how to interact with functionality (controls, forms or functionality)***

Frequency: 17 instances (1.82%)

Median severity rating: 3 (major)

User problem

Users had no indication of how to interact with functionality. In one example, users were presented with a sliding bar to select the budget to filter cars they could choose. However, before clicking on the control for the first time, there was no indication that it would work as a sliding bar that could be dragged with a mouse, as shown in Figure D.10. In another example, users had to pick a seat in a seat map. The interface showed small circles to identify the seats, but there was no indication that users had to click on a circle to pick a seat at that position.

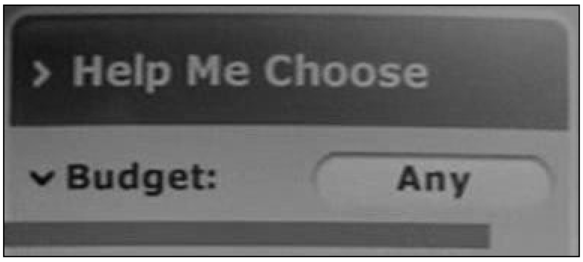


Figure D.10 – Example of control with no indication of how to interact with

Technical cause

The main cause of problems in this subcategory was the use of interactive components that did not have an obvious behaviour to users. In the case of the sliding bar on the Ford website, for example, it did not use interface components that users were already used to.

- ***Subcategory: No/insufficient feedback to inform that action has had an effect (controls, forms or functionality)***

Frequency: 15 instances (1.6%)

Median severity rating: 3 (major)

User problem

Users had problems when they had no or insufficient feedback about the completion of their actions. In one example, users wanted to list cars that matched some requirements. After selecting their requirements, the only feedback given was the change in the colour of rings around cars. This was not sufficient for users to perceive that their action of selecting the requirements had had any effect. Figure D.11 illustrates the change in colour, with the car in the centre having a slightly darker ring around it to indicate that it matches the requirements.

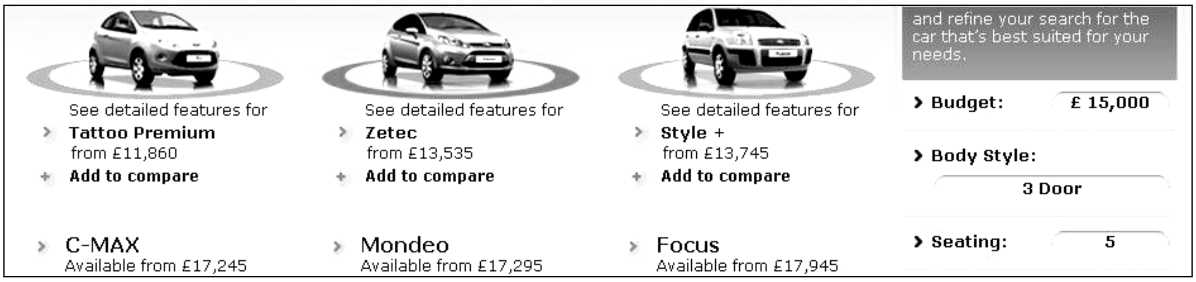


Figure D.11 – Example of insufficient feedback of completion of actions – feedback was given only by changing colour of rings around cars that matched requirements

Technical causes

In the example shown on the Ford website, the main issue was the sole use of colour to provide feedback for actions. Other issues also included the use of unclear messages and the placement of feedback messages in unobvious places that were out of users' viewport when they were using screen magnifiers or had to use scroll bars due to the increased font size.

- ***Subcategory: It is not clear what particular controls or form elements do(controls, forms or functionality)***

Frequency: 14 instances (1.5%)

Median severity rating: 3 (major)

User problem

Users could not know what particular elements did, due to poor affordance. Most examples with partially sighted users involved cases where users encountered an icon or another graphical element with poor affordance.

Technical causes

Icons or graphical elements do not clearly indicate what they do to users. Examples include icons that do not convey any meaning of the feature they represent.

- ***Subcategory: No obvious way to return to homepage (navigation)***

Frequency: 13 instances (1.39%)

Median severity rating: 3 (major)

User problem

Users found it difficult to return to the home page of a website due to not having a link to go to the home page, or to having a link to the home page that was not obvious.

Technical cause

In some websites, the image of a logo was not made into a link to the home page, as many users expected. In other websites, users expected that there would be a link named “home”, but the websites used other names that were not obvious for users to identify.

- ***Subcategory: No/insufficient feedback that users have not performed a required action (controls, forms or functionality)***

Frequency: 12 instances (1.28%)

Median severity rating: 3 (major)

User problem

Users encountered problems when they had no or insufficient feedback that they had not performed a required action. In one example, users selected a seat for an event next to the last seat in a row, receiving a message “you cannot leave an unoccupied seat”. They had no idea of what to do next after receiving this message. In other examples, information about empty required fields was not highlighted enough.

Technical causes

Unclear messages and poorly presented messages were the main causes of problems in this subcategory. Some messages were placed that were not in users’ viewport or not positioned in places that were not obvious for users to find.

- ***Subcategory: Information architecture too complex (information architecture)***

Frequency: 12 instances (1.28%)

Median severity rating: 3 (major)

User problem

Users reported problems when the information architecture in websites were too complex, making them have to follow too many steps in the navigation to arrive at the content they wanted.

Technical cause

The cause of problems in this subcategory was having too many depths in the navigation structure of a website. This made users have to follow too many links (or navigation steps) to arrive at the content they needed.

- ***Subcategory: Users cannot understand sequence of interaction (controls, forms of functionality)***

Frequency: 10 instances (1.07%)

Median severity rating: 3 (major)

User problem

Like blind users, partially sighted users also encountered problems when a task involved a sequence of actions or steps had to be performed in a certain order, but the order was not clear to them. The most common issue encountered by partially sighted users in problems in this subcategory was also with the Ford website, where users tried to compare different cars available on a website and tried to use the "compare" button. However, they could not understand that they had to select up to three cars to compare before asking to compare them.

Technical causes

The interaction model of interfaces where those problems occurred were not logical to users and did not have clear indications of the step-by-step process that should be taken to complete tasks.

- ***Subcategory: Link grouping poor (links)***

Frequency: 9 instances (0.96%)

Median severity rating: 3 (major)

User problem

Users found it difficult to find a specific link when they were not properly grouped. They found that some groups of links did not belong to the same category as they would expect, or find long lists of links with no logical organisation.

Technical causes

The main cause of problems in this subcategory was the poor organisation of groups of links. In a list of sandwiches, for example, users expected that related sandwiches would be grouped together. In other cases, it did not make sense to users to have unrelated links at the same group.

- ***Subcategory: Moving multimedia content is distracting (audio, video or multimedia)***

Frequency: 7 instances (0.75%)

Median severity rating: 3 (major)

User problem

Users reported problems with animations placed on web pages, as it caused distractions to them and made it harder to concentrate on the content they had to read to accomplish their tasks.

Technical causes

Distracting animations was placed next to important content, with no option to stop or pause.

- ***Subcategory: No alternative to text in specific format (text)***

Frequency: 5 instances (0.75%)

Median severity rating: 3 (major)

User problem

Users reported problems when there was no alternative to text in PDF. They found it considerably harder to use the features available in their assistive technology to personalise features such as font size and colour. With font size, for example, unless they changed the size on a PDF reader, the resizing operation done by the screen magnifier made the text become pixelated.

Technical causes

No alternative to text available on PDF only, that did not offer the same compatibility with assistive technologies as plain text in HTML.

- ***Subcategory: Graphic or multimedia not compatible with assistive technology (underlying system characteristics)***

Frequency: 6 instances (0.64%)

Median severity rating: 3 (major)

User problem

Users could not watch a video embedded in a website with their screen magnifier. They could only see a black square on the screen in the place where the video was meant to be shown.

Technical causes

Embedded videos in Windows media format were not compatible with the screen magnifiers ZoomText and Supernova.

- *Subcategory: Navigation bar not salient (navigation)*

Frequency: 6 instances (0.64%)

Median severity rating: 3 (major)

User problem

Users found it difficult to identify where the navigation bar was located in some websites. They found that the navigation bar was not salient enough. In some cases, users found it difficult to locate the navigation bar when it was not located at the main viewport that was open when they arrived at a page.

Technical causes

Poor layout of navigation bars was the main cause of problems in this subcategory. Navigation bars that were not salient did not have good colour contrast and font size that would make them stand out from the rest of the page and make them easier to find visually for users. In some cases, the navigation bar was located at the right-hand side of a website, which was often not shown immediately when users with a limited viewport opened the website.

- *Subcategory: No alternative to information presented in tables (table)*

Frequency: 5 instances (0.53%)

Median severity rating: 3 (major)

User problem

Some users found reading information on tables extremely difficult. In some cases, the level of magnification in their assistive technologies made it very complicated and laborious for users to relate different columns of tables, due to the need to pan across different parts of the screen to be able to read the tables in a small viewport. In those cases, not having an alternative to information laid-out in tables made it very difficult for users to complete their tasks.

Technical causes

For problems in this subcategory, the main technical cause was not having an alternative to tables. Users who encountered problems with tables had severe difficulties to process information laid-out in tables, and would need an alternative.

- ***Subcategory: Inconsistent navigation structure in different pages (navigation)***

Frequency: 5 instances (0.53%)

Median severity rating: 3 (major)

User problem

Users encountered problems when the place of navigation structures was inconsistent in different pages. For users with screen magnifiers, this meant that they would have to spend longer to find where a navigation bar was, even though they had already learnt where the navigation bar was before.

Technical causes

Having the navigation bar at different places and with different layout were the main cause of the inconsistencies reported in this subcategory.

- ***Subcategory: Meaning in content is lost or modified due to transformation (content)***

Frequency: 4 instances (0.43%)

Median severity rating: 3 (major)

User problem

Users needed to find specific information about a legal agreement on a law office website. A simplified version of the agreement was provided, but users reported problems as they were not provided with the information they needed from the simplified version.

Technical causes

Simplification of text did not include all relevant information provided in full version of text.

- *Subcategory: Default presentation of table not adequate (table)*

Frequency: 4 instances (0.43%)

Median severity rating: 3 (major)

User problem

Users reported problems when the presentation of tables made it harder for them to read information in them. In one example, users were seeking nutritional information of a specific sandwich. However, they found it difficult to relate specific numeric information to related labels due to poor presentation. Figure D.12 shows the table where the problems occurred, with no visual cues to help users visually relate labels and figures.

Energy Values (kcal) per pack	545
Protein (g) per pack	30.7
Carbohydrate (g) per pack	40
– of which sugars(g)	5.3
Fat (g) per pack	29
– of which saturates (g)	5.1
Dietary Fibre (g) per pack	4
Sodium (g) per pack	0.878
Average weight (g) per pack	280

Figure D.12 – Example of inadequate presentation of table – no visual cues to relate labels and figures of nutritional information

Technical problems

Poor layout of tables was the main cause of problems in this subcategory. The main issues were related to columns that were too far from each other or the lack of borders or visual cues to help relate labels and figures.

- *Subcategory: Multimedia starting automatically is irritating (audio, video or multimedia)*

Frequency: 4 instances (0.43%)

Median severity rating: 3 (major)

User problem

Users became irritated when multimedia started automatically, especially when it had sound. It took users some time until they could find where on the screen a pause or stop button was available for them, particularly with high levels of magnification.

Technical causes

Starting multimedia content without users' request was the main cause of problems in this subcategory.

- ***Subcategory: Language too complicated for perceived target audience (content)***

Frequency: 3 instances (0.43%)

Median severity rating: 3 (major)

User problem

Like blind users, partially sighted users also found it difficult to understand the language level in pages related to their tasks, including difficult terms or difficult grammatical structures. In a legal website targeted at disabled people, users were seeking an explanation of the term "structured negotiation". One would expect that the FAQ (Frequently Asked Questions) would not be targeted at experts, but users encountered an explanation with terms they found difficult, such as "advocacy" or "litigation".

Technical causes

Use of difficult language without proper alternatives or explanations in a difficulty level that could be understood by non-experts in the topic of the website.

3. Description of the main problems encountered by dyslexic users

This section presents a description of the main problems encountered by dyslexic users, focusing on the problems that occurred most frequently and on the most severe problems. Firstly, the problems that were listed both as most frequent and most severe are presented. The section proceeds with the presentation of the problems listed as most severe only and the problems listed as most frequent but with low severity.

3.1. Problems encountered by dyslexic users with high frequency and high median severity

- ***Subcategory: Content not found in pages where expected by users (content)***

Frequency: 112 instances (16.16%)

Median severity rating: 3 (major)

User problem

Like blind and partially sighted users, dyslexic users also encountered problems when they could not find content in pages where they expected. Users confidently followed a link to a page, but a piece of information that they expected to find there was missing. For example, on a restaurant website, users sought information about the price of a platter for delivery, and found a page with a description of platters. They expected that the page would list all information about platter, including prices, but they could not find any information about prices.

Technical causes

The main technical cause of problems in this subcategory was connected to the information architecture and organisation of websites. Users expected to find certain pieces of information in parts of the website that were different from that laid-out by the designers of the websites.

- ***Subcategory: Navigation elements do not help users find what they are seeking (navigation)***

Frequency: 87 instances (12.55%)

Median severity rating: 3 (major)

User problem

Like blind and partially sighted users, dyslexic users also found problems with navigation elements that were confusing and disorienting, and did not help them find the information they were seeking in their task. In one example of problem in this subcategory, users were seeking tips for driving tests targeted at young people. The navigation several options that seemed to be plausible, such as “Motoring Advice” and “Driving School””, but users could not be sure which one to follow. Figure D.13 shows the navigation bar of The Automobile Association website with the options available.

Figure D.13 – Example of navigation of The Automobile Association website – users had difficulties finding where to find information about driving tests for young people

Technical causes

In most cases when problems in this subcategory occurred, there were not specific problems with individual links available in the navigation. The reason for the problems was with the overall structure of the navigational elements, caused by a poorly designed information architecture.

- ***Subcategory: Expected functionality not present (controls, forms or functionality)***

Frequency: 37 instances (5.34%)

Median severity rating: 3 (major)

User problem

Users reported problems when they expected websites to have a functionality that was not present. The most frequent problems reported by users were the lack of search features and features of auto-completion in form fields.

In the case of the search features, many users reported that they prefer to use a search when using websites to speed up their navigation. When a search feature was not available, they reported that this prevented them from using a navigation strategy that they were used to.

Users with spelling difficulties reported that they would benefit significantly from having an auto-completion feature in form fields. With this feature, they could type the first few letters of the word they wanted, and the system would provide them with suggestions of words that begin with the letters they typed. Not having this feature made it more difficult for users to spell words they had to type in form fields.

Technical causes

The main technical causes of problems in this subcategory were not providing a search feature to users and not implementing a feature to auto-complete forms. Auto-completion features use a database of words and suggest to users a list of words or phrases that begin with the letters informed by users.

- ***Subcategory: Users cannot make sense of content (content)***

Frequency: 29 instances (4.18%)

Median severity rating: 3 (major)

User problem

Like blind and partially sighted users, dyslexic users also encountered problems when they could not make sense of content in websites. Users often encountered incomplete information or unclear explanations about content that was specific to an unfamiliar domain in a website. On the TicketMaster website, users were trying to buy tickets for an event. They encountered a seat plan with circles with different colours that represented colours. Available seats were orange and unavailable seats were grey. Users tried to click on several different seats using trial and error to figure out which seats were available, since they could not infer this just by the colours that were used.

Technical causes

Incomplete and unclear information and explanations were the main causes of problems in this subcategory. Many websites used colours or names that were not familiar to users to convey information, making it difficult for users to make sense of the content they saw on the screen.

- ***Subcategory: Functionality does not work (as expected) (controls, forms and functionality)***

Frequency: 29 instances (4.18%)

Median severity rating: 3 (major)

User problem

Like blind and partially sighted users, dyslexic users also reported problems when functionality did not work as they expected. The most notorious of those problems with dyslexic users was with search features that did not return any results with keywords that users expected would be in the website. Other problems were related to functionality that simply did not respond to users' actions or behaved in an unexpected manner.

Technical causes

Faulty implementation of features was the main cause of problems in this subcategory. In the particular case of search features, many websites included only a fraction of web pages in the index of the search, not including static pages, for example.

- ***Subcategory: Users cannot understand sequence of interaction (controls, forms or functionality)***

Frequency: 17 instances (2.45%)

Median severity rating: 3 (major)

User problem

Like blind and partially sighted users, dyslexic users also encountered problems when a task involved a sequence of actions or steps had to be performed in a certain order, but the order was not clear to them. Besides the issues with the Ford website that were also encountered by partially sighted users, dyslexic users also encountered many problems with the York City Council website. Users were seeking local services based on a given address. After typing an address, users had to select an address from a list before proceeding (in case there were several flats under the same house number, for example). However, other links, such as show services on map were available on the screen, and users did not know they had to select the address before viewing the services. Figure D.14 illustrates the screen where those options were shown.

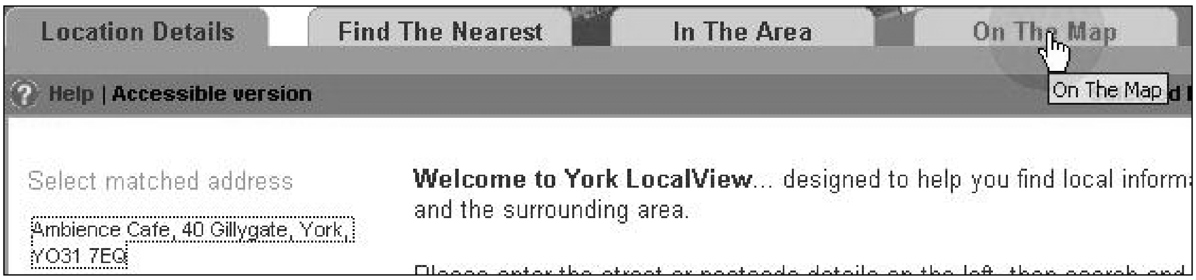


Figure D.14 – Example of problem where users could not identify sequence of interaction on the York City Council website – an address had to be selected before viewing services on a map

Technical causes

The interaction model of interfaces where those problems occurred were not logical to users and did not have clear indications of the step-by-step process that should be taken to complete tasks.

3.2. Problems encountered by dyslexic users with high median severity

- ***Subcategory: No/insufficient feedback that users have not performed a required action (controls, forms or functionality)***

Frequency: 7 instances (1.01%)

Median severity rating: 3 (major)

User problem

Like blind and partially sighted users, dyslexic users also encountered problems when they were given no or insufficient feedback that they had not performed a required action. In one example at the York City Council website, users were seeking local services in their area. After they typed the house numbers, street name and post code, users would have to select an address from a list (in case there were more flats under the same number, for example). However, the house number given had only one address related to it. Users only found a very faint message “select address”, that was not sufficient to help them identify that they had to select an address.

Technical causes

Unclear messages and poorly presented messages were the main causes of problems in this subcategory.

- ***Subcategory: Navigation elements not understandable (e.g. using jargon of difficult language) (navigation)***

Frequency: 5 instances (0.72%)

Median severity rating: 3 (major)

User problem

Like blind users and partially sighted users, dyslexic users also find problems when they could not understand the meaning of jargon and difficult words in the navigation of websites.

Technical cause

The cause of those problems was the use of jargon and difficult words as navigation elements.

3.3. Problems encountered by dyslexic users with high frequency and lower median severity

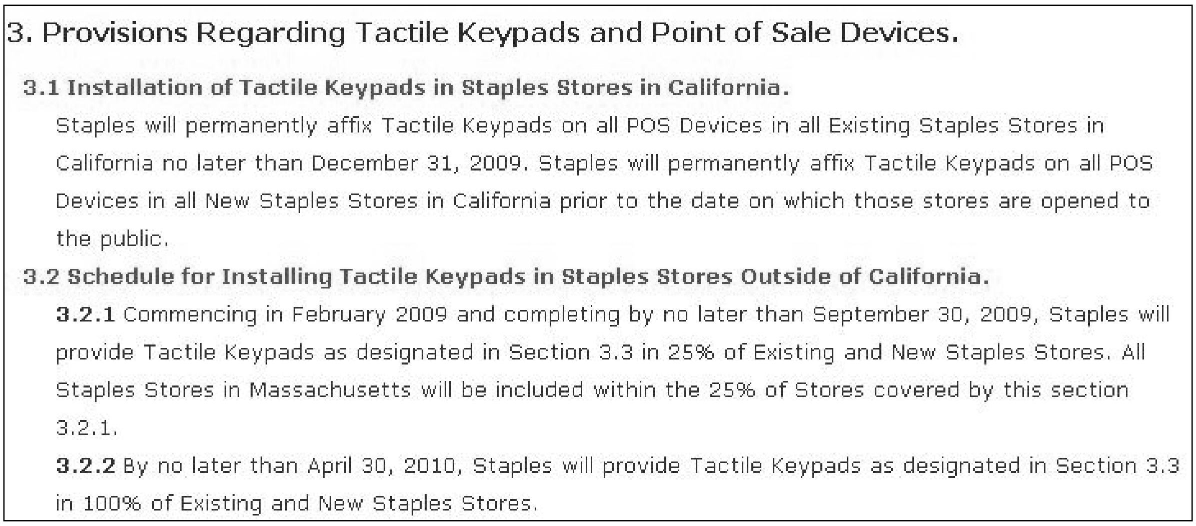
- ***Subcategory: Difficult to scan page for specific item (all media types)***

Frequency: 72 instances (10.39%)

Median severity rating: 2 (minor)

User problem

This subcategory is related to problems when the user encounters difficulties scanning for specific items in a web page. Users could not find visual aids that would make the content they needed stand out from the rest of the web page. In the Lflegal website, users were seeking the deadline for the installation of tactile keypads in a store chain outside of California. As shown in Figure D.15, the date was contained in the middle of the paragraph 3.2.1, which was part of a very long document. Users found it difficult to scan the text with no visual cues to that date.



3. Provisions Regarding Tactile Keypads and Point of Sale Devices.

3.1 Installation of Tactile Keypads in Staples Stores in California.

Staples will permanently affix Tactile Keypads on all POS Devices in all Existing Staples Stores in California no later than December 31, 2009. Staples will permanently affix Tactile Keypads on all POS Devices in all New Staples Stores in California prior to the date on which those stores are opened to the public.

3.2 Schedule for Installing Tactile Keypads in Staples Stores Outside of California.

3.2.1 Commencing in February 2009 and completing by no later than September 30, 2009, Staples will provide Tactile Keypads as designated in Section 3.3 in 25% of Existing and New Staples Stores. All Staples Stores in Massachusetts will be included within the 25% of Stores covered by this section 3.2.1.

3.2.2 By no later than April 30, 2010, Staples will provide Tactile Keypads as designated in Section 3.3 in 100% of Existing and New Staples Stores.

Figure D.15 – Example of from Lflegal – users found it difficult to scan for part of text in the middle of paragraph with no visual cues

Technical causes

The cause of problems in this subcategory was the lack of visual cues to help users scan for specific content on pages, be it text, images or other visual elements. The lack of a clear visual structure and highlighting of specific parts made it harder for dyslexic users to scan for specific items without having to actually read through long pages.

- ***Subcategory: Default presentation of text not adequate (text)***

Frequency: 44 instances (6.35%)

Median severity rating: 2 (minor)

User problem

Users reported difficulties to read text with specific formats. Common issues reported by users included difficulties reading text using italics, with inadequate spacing between lines and paragraphs, small font size, inappropriate font style, and inappropriate colour background.

Many users encountered problems with black writing on white background. For these users, reading text on white background for a long time caused the text to start forming “visual patterns”, or “dancing around”. Although most web browsers have features to change the colour background of a website, none of the participants of this study knew about this feature, or if they knew, they found it very difficult to use. In most cases when this problem was reported, participants expected that the websites would provide them with a colour selector feature instead.

Technical causes

Several issues related to the presentation of text caused problems in this subcategory, including font size, typeface, line spacing and text background and foreground colours.

Regarding background and foreground colours, the cause of the problems was the lack of a feature to enable users to select their preferred colours embedded in websites. None of the users used browser settings to change colours.

Typeface fonts that make it difficult to use also caused problems to users. This included font styles that were difficult to read, such as fonts with serif, and use of italics.

Small font size was also a very frequent cause of problems to users. None of the users used any size adjustments on their web browsers, but reported that having a default font size too small was a problem for them.

- ***Subcategory: Too much information on page (content)***

Frequency: 34 instances (4.91%)

Median severity rating: 2 (minor)

User problem

Users became disoriented and overwhelmed when they pages had too much information on them. Dyslexic users found it harder to focus their attention and to look for information when pages were crowded with too much information.

Technical causes

The cause of problems in this subcategory was the excessive amount of information that was presented all at once in one web page. Those pages would often lack structural elements to break their parts into smaller blocks that would be easier for users to read.

- ***Subcategory: Organisation of content is inconsistent with web conventions/common sense (content)***

Frequency: 30 instances (4.33%)

Median severity rating: 2 (minor)

User problem

Users encountered problems when the organisation of information was not consistent to conventions they were used to or with common sense. In one example, on an education research website, users found several links scattered in a squared area with no logical arrangement. This made it very difficult for users to read through the list of links. Figure D.16 shows the box containing those links from the JISC website.

Figure D.16 – Example of from JISC – content on page was not logically organised

Technical causes

The problems with illogical organisation included issues with related information not being displayed logically along with other related information, illogical ordering of information (not in alphabetical order, for example) and lack of patterns in the way information is listed.

- ***Subcategory: No/insufficient feedback to inform that action has had an effect (controls, forms or functionality)***

Frequency: 17 instances (2.45%)

Median severity rating: 2 (minor)

User problem

Users reported problems when they had no or insufficient feedback to inform an action has had effect. In one example, users did not realise that a selection of requirements to list cars had taken effect, as the feedback was only given by changes in the colour of rings around cars that did not match the requirements. In another example in the York City Council website, users entered their address information to search for local services, and the following screen only asked them to select an address, which was not enough to signal to users that their previous action had had an effect.

Technical causes

The lack of clear indication of feedback was the main cause of problems in this subcategory. Many problems were related to feedback that was too subtle for users to perceive or unclear messages that did not fully inform users about the status of the system after they performed an action.

- ***Subcategory: Link destination not clear (link)***

Frequency: 16 instances (2.31%)

Median severity rating: 2 (minor)

User problem

Users encountered problems when the destination of certain links was not clear. For example, on the Trades Union Congress (TUC) website, users were seeking information about a Teachers' Union. On a page that listed affiliated unions, they encountered a link "Education and Training", and were not sure if it was a link for unions related to education or a link to general information about education and training.

Technical causes

Unclear description of links was the cause of problems in this subcategory. In some cases, such as illustrated in the TUC example, a link taken out of context could make sense to users, but when the link was placed in the context of different pages, its destination could have different meanings to users.

- ***Subcategory: Language too complicated for perceived target audience (content)***

Frequency: 15 instances (2.16%)

Median severity rating: 2 (minor)

User problem

Like blind and partially sighted users, dyslexic users also found it difficult to understand the language level in pages related to their tasks, including difficult terms or difficult grammatical structures. In a legal website targeted at disabled people, users were seeking an explanation of the term “structured negotiation”. One would expect that the FAQ (Frequently Asked Questions) would not be targeted at experts, but users encountered an explanation with terms they found difficult, such as “advocacy” or “litigation”.

Technical causes

Use of difficult language without proper alternatives or explanations in a difficulty level that could be understood by non-experts in the topic of the website.

- ***Subcategory: Irrelevant content before task content (content)***

Frequency: 13 instances (1.88%)

Median severity rating: 2 (minor)

User problem

Dyslexic users also reported problems when they found irrelevant content before content that was relevant to their tasks. Irrelevant search results not related to their search were shown before relevant results. In other examples, advertisement information in a car insurance web page was shown before information about insurance plans.

Technical causes

The order of content in the page did not follow the logical order that would be expected by users, with the most relevant content appearing first. In the case of search features, the cause was due to poor indexing algorithms in search engines.

- ***Subcategory: Users inferred the existence of functionality where there was not one (controls, forms or functionality)***

Frequency: 13 instances (1.88%)

Median severity rating: 2 (minor)

User problem

Dyslexic users also encountered problems with parts of the page that looked like they had some functionality, mainly statements in the imperative form, such as “Book a redelivery”. This made users believe that the specific piece of text would work as a link or a button, but when they tried to press Enter on the text, it did not act as an activator of any functionality.

Technical cause

Text with statements in the imperative form that looked like activators of some functionality were not turned into a link or a button, as users would expect.

References

Abascal, J., Arrue, M., Fajardo, I., Garay, N. and Tomás, J. (2004) 'The use of guidelines to automatically verify Web accessibility', *Univers. Access Inf. Soc.,* 3(1), 71-79.

Al-Khalifa, H. (2012) 'The accessibility of Saudi Arabia government Web sites: an exploratory study', *Universal Access in the Information Society,* 11(2), 201-210.

Al-Wabil, A., Zaphiris, P. and Wilson, S. (2006) 'Web Design for Dyslexics: Accessibility of Arabic Content' in Miesenberger, K., Klaus, J., Zagler, W. and Karshmer, A., eds., *Computers Helping People with Special Needs*, Springer Berlin / Heidelberg, 817-822.

Al-Wabil, A., Zaphiris, P. and Wilson, S. (2007) 'Web Navigation for Individuals with Dyslexia: An Exploratory Study' in Stephanidis, C., ed. *Universal Acess in Human Computer Interaction. Coping with Diversity*, Springer Berlin / Heidelberg, 593-602.

Alonso, F., Fuertes, J. L., González, Á. L. and Martínez, L. (2010) 'On the testability of WCAG 2.0 for beginners', in *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*, Raleigh, North Carolina, ACM, 1-9.

Andre, T. S., Rex Hartson, H., Belz, S. M. and McCreary, F. A. (2001) 'The user action framework: a reliable foundation for usability engineering support tools', *International Journal of Human-Computer Studies,* 54(1), 107-136.

Arditi, A. and Faye, E. (2004) 'Monocular and Binocular Letter Contrast Sensitivity and Letter Acuity in a Diverse Ophthalmologic Practice', *Supplement to Optometry and Vision Science,* 12S(81).

Arrue, M., Fajardo, I., Lopez, J. and Vigo, M. (2007) 'Interdependence between technical web accessibility and usability: its influence on web quality models', *International Journal of Web Engineering and Technology,* 3(3), 307-328.

Australian Government (2010) 'Web Guide: Accessibility', [online], available: http://webguide.gov.au/accessibility-usability/accessibility/ [accessed 21/02/2013].

Babu, R. and Singh, R. (2009) 'Evaluation of Web Accessibility and Usability from Blind User's Perspective: The Context of Online Assessment', in (AISeL), A. E. L., ed.

Americas Conference on Information Systems (AMCIS) 2009, San Francisco, Paper 623.

Benavídez, C., Fuertes, J. L., Gutiérrez, E. and Martínez, L. (2006) 'Teaching Web Accessibility with Contramano and Hera', in *Proceedings of 10h Internacional Conference on Computers Helping People With Special Needs*, Springer Berlin / Heidelberg, 341-348.

Bradford, J. (2005) 'Designing Web pages for dyslexic users', *Dyslexia Online Magazine* [online], available: http://www.dyslexia-parent.com/mag35.html [accessed 29/06/2011].

Brajnik, G. (2004) 'Comparing accessibility evaluation tools: a method for tool effectiveness', *Univers. Access Inf. Soc.,* 3, 252-263.

Brajnik, G. (2006) 'Web Accessibility Testing: When the Method Is the Culprit' in Miesenberger, K., Klaus, J., Zagler, W. and Karshmer, A., eds., *Computers Helping People with Special Needs*, Springer Berlin / Heidelberg, 156-163.

Brajnik, G. (2008) *A comparative test of web accessibility evaluation methods,* translated by New York, NY, USA: ACM, 113-120.

Brajnik, G., Yesilada, Y. and Harper, S. (2010) *Testability and validity of WCAG 2.0: the expertise effect,* translated by New York, NY, USA: ACM, 43-50.

Brazilian Government (2011) '*e-MAG: Accessibility Model for e-Government*', [online], available: http://emag.governoeletronico.gov.br/emag/ [accessed 21/02/2013].

Brewer, J. (2004) 'W3C Web Accessibility Initiative Statement on Web Access Report from UK Disability Rights Commission', *Web Accessibility Initiative (WAI), the World Wide Web Consortium (W3C)* [online], available: http://www.w3.org/2004/04/wai-drc-statement.html [accessed 09/11/2011].

British Dyslexia Association (2011) 'Dyslexia Style Guide', [online], available: http://www.bdadyslexia.org.uk/about-dyslexia/further-information/dyslexia-style-guide.html [accessed 29/06/2011].

British Standards Institute (2010) 'BS 8878:2010 - Web accessibility Code of practice', 90.

Caldwell, B., Cooper, M., Reid, L. G. and Vanderheiden, G. (2008) 'Web Content Accessibility Guidelines 2.0', *Web Accessibility Initiative (WAI), World Wide Web Consortium (W3C)* [online], available: http://www.w3.org/TR/WCAG20 [accessed 14/04/2011].

Chisholm, W., Vanderheiden, G. and Jacobs, I. (1999) 'Web Content Accessibility Guidelines 1.0', *Web Accessibility Initiative (WAI), World Wide Web Consortium (W3C)* [online], available: http://www.w3.org/TR/WCAG10/ [accessed 14/04/2011].

Chisholm, W., Vanderheiden, G. and Jacobs, I. (2000) 'Techniques for Web Content Accessibility Guidelines 1.0 ', [online], available: http://www.w3.org/TR/WAI-WEBCONTENT-TECHS/ [accessed 25/01/2012].

Choi, Y. S., Yi, J. S., Law, C. M. and Jacko, J. A. (2006) 'Are "universal design resources" designed for designers?', in *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*, New York, NY, USA, ACM, 87-94.

Colwell, C. and Petrie, H. (2001) 'Evaluation of guidelines for designing accessible Web content', *SIGCAPH Comput. Phys. Handicap.,* (70), 11-13.

Cooper, M., Guarino Reid, L., Vanderheiden, G. and Caldwell, B. (2010a) '*Understanding WCAG 2.0: A guide to understanding and implementing Web Content Accessibility Guidelines 2.0*.', *Web Accessibility Initiative (WAI), World Wide Web Consortium (W3C)* [online], available: http://www.w3.org/TR/UNDERSTANDING-WCAG20/ [accessed 28/11/2011].

Cooper, M., Reid, L. G., Vanderheiden, G. and Caldwell, B. (2010b) 'Techniques and Failures for Web Content Accessibility Guidelines 2.0', [online], available: http://www.w3.org/TR/2010/NOTE-WCAG20-TECHS-20101014/ [accessed 17/01/2012].

Cooper, M., Sloan, D., Kelly, B. and Lewthwaite, S. (2012) 'A challenge to web accessibility metrics and guidelines: putting people and processes first', in *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility*, Lyon, France, 2207028: ACM, 1-4.

Coyne, K. P. and Nielsen, J. (2001) *Beyond ALT Text: Making the Web Easy to Use for Users with Disabilities*

Craig, J. and Cooper, M. (2011) 'Accessible Rich Internet Applications (WAI-ARIA) 1.0 ', [online], available: http://www.w3.org/TR/wai-aria/ [accessed

Disability Rights Commission (2004) *The Web: access and inclusion for disabled people: A formal Investigation conducted by the Disability Rights Commission*, London: The Stationery Office.

Donnelly, A. and Magennis, M. (2003) 'Making accessibility guidelines usable', in *Proceedings of the User interfaces for all 7th international conference on Universal access: theoretical perspectives, practice, and experience*, Paris, France, 1765431: Springer-Verlag, 56-67.

Ericsson, K. A. and Simon, H. A. (1993) *Protocol Analysis - Rev'd Edition: Verbal Reports as Data,* A Bradford Book.

Espadinha, C., Pereira, L. M., da Silva, F. M. and Lopes, J. B. (2011) 'Accessibility of Portuguese Public Universities' sites', *Disability and Rehabilitation,* 33(6), 475-485.

European Commission (2012) 'The EU Internet Handbook: Accessibility', [online], available: http://ec.europa.eu/ipg/standards/accessibility/index_en.htm [accessed 21/02/2013].

Evans-Cowley, J. S. (2005) 'The Accessibility of Municipal Government Websites', *Journal of EGovernment,* 2(2), 75-90.

Evett, L. and Brown, D. (2005) 'Text formats and web design for visually impaired and dyslexic readers--Clear Text for All', *Interacting with Computers,* 17(4), 453-472.

Friedman, M. G. and Bryen, D. N. (2007) 'Web accessibility design recommendations for people with cognitive disabilities', *Technology and Disability,* 19(4), 205-212.

Goette, T., Collier, C. and White, J. (2006) 'An exploratory study of the accessibility of state government Web sites', *Universal Access in the Information Society,* 5(1), 41-50.

Hackett, S. and Parmanto, B. (2005) 'A longitudinal evaluation of accessibility: higher education web sites', *Internet Research,* 15(3), 281 - 294.

Hackett, S., Parmanto, B. and Zeng, X. (2005) 'A retrospective look at website accessibility over time', *Behaviour & Information Technology,* 24(6), 407 - 417.

Hanson, V. L. (2001) 'Web access for elderly citizens', in *Proceedings of the 2001 EC/NSF workshop on Universal accessibility of ubiquitous computing: providing for the elderly*, Alc\ácer do Sal, Portugal, 564531: ACM, 14-18.

Hanson, V. L. (2009) 'Age and web access: the next generation', in *Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibililty (W4A)*, Madrid, Spain, 1535658: ACM, 7-15.

Hanson, V. L., Richards, J. T., Harper, S. and Trewin, S. (2009) 'Web accessibility' in Stephanidis, C., ed. *The Universal Access Handbook*, Boca Raton, FL: CRC Press, 10-1 - 10-15.

Harrison, C. (2008) *Exploring Emotional Web Experience: More than just usability and good design*, unpublished thesis University of York.

Harrison, C. and Petrie, H. (2007) 'Severity of Usability and Accessibility Problems in eCommerce and eGovernment Websites' in Bryan-Kinns, N., Blanford, A., Curzon, P. and Nigay, L., eds., *People and Computers XX — Engage*, Springer London, 255-262.

Henry, S. L. (2003) 'Chapter 1: Understanding Web Accessibility' in Thatcher, J., C.D., W., S.L., H., S., S., M.D., U., M., B., B., R. and Bohman, P., eds., *Constructing Accessible Websites*, Birmingham: Glasshaus.

Hornbæk, K. and Frøkjær, E. (2008) 'Comparison of techniques for matching of usability problem descriptions', *Interact. Comput.,* 20(6), 505-514.

Hégaret, P. L. and Smeman, S. (2012) 'CSS Validation service', [online], available: http://jigsaw.w3.org/css-validator/ [accessed 06/01/2012].

IBM (2011) 'Rational Policy Tester Accessibility Edition', [online], available: http://www-01.ibm.com/software/awdtools/tester/policy/accessibility/ [accessed 09/11/2011].

International Standards Organization (1993) 'ISO 9241-3, Ergonomic requirements for office work with visual display terminals (VDTs) - Part 3: Visual display requirements. Amendment 1.',

International Standards Organization (1998) 'ISO 9241-11 - Ergonomic requirements for office work with visual display terminals (VDTs)-Part 11: guidance on usability—Part 11: guidance on usability',

International Standards Organization (2008) 'ISO 9241-171: Ergonomics of human-system interaction. Part 171: Guidance on software accessibility',

Ivory, M. Y. (2003) *Automated Web Site Evaluation : Researchers' and Practitioners' Perspectives,* Kluwer Academic Publishers.

Jacobs, I., Gunderson, J. and Hansen, E. (2002) 'User Agent Accessibility Guidelines 1.0', *Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C)* [online], available: http://www.w3.org/TR/UAAG10/ [accessed 09/11/2011].

Juicy Studio (2012) 'Colour Contrast Analyser ', [online], available: http://juicystudio.com/services/aertcolourcontrast.php [accessed 06/01/2012].

Kane, S. K., Shulman, J. A., Shockley, T. J. and Ladner, R. E. (2007) 'A web accessibility report card for top international university web sites', in *Proceedings of the 2007 international cross-disciplinary conference on Web accessibility (W4A)*, Banff, Canada, 1243472: ACM, 148-156.

Kelly, B. (2002) 'An Accessibility Analysis of UK University Entry Points', *Ariadne,* (33).

Kelly, B., Sloan, D., Brown, S., Seale, J., Petrie, H., Lauke, P. and Ball, S. (2007) 'Accessibility 2.0: people, policies and processes', in *Proceedings of the 2007 international cross-disciplinary conference on Web accessibility (W4A)*, Banff, Canada, 1243471: ACM, 138-147.

Kelly, B., Sloan, D., Phipps, L., Petrie, H. and Hamilton, F. (2005) 'Forcing standardization or accommodating diversity?: a framework for applying the WCAG in the real world', in *Proceedings of the 2005 International Cross-Disciplinary Workshop on Web Accessibility (W4A)*, Chiba, Japan, 1061820: ACM, 46-54.

Kolatch, E. (2000) 'Designing for users with cognitive disabilities', *The Universal Usability Guide, University of Maryland, College Park* [online], available: http://www.otal.umd.edu/UUGuide/erica [accessed 29/06/2011].

Koyani, S., Bailey, R., Nall, J., Allison, S., Mulligan, C., Bailey, K. and Tolson, M. (2004) 'Research-based Web design and usability guidelines.', [online], available: http://www.usability.gov/guidelines/guidelines_book.pdf [accessed 14/11/2011].

Kurniawan, S. and Zaphiris, P. (2005) 'Research-derived web design guidelines for older people', in *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility*, Baltimore, MD, USA, 1090810: ACM, 129-135.

Law, E. C., Hvannberg, E., Cockton, G., Palanque, P., Scapin, D., Springett, M., Stary, C. and Vanderdonckt, J. (2005) 'Towards the Maturation of IT Usability Evaluation (MAUSE)' in Costabile, M. and Paternò, F., eds., *Human-Computer Interaction - INTERACT 2005*, Springer Berlin Heidelberg, 1134-1137.

Law, E. L.-C. and Hvannberg, E. T. (2008) 'Consolidating usability problems with novice evaluators', in *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*, Lund, Sweden, ACM,

Lazar, J., Allen, A., Kleinman, J. and Malarkey, C. (2007) 'What Frustrates Screen Reader Users on the Web: A Study of 100 Blind Users', *International Journal of Human-Computer Interaction,* 22(3), 247-269.

Lazar, J., Beavan, P., Brown, J., Coffey, D., Nolf, B., Poole, R., Turk, R., Waith, V., Wall, T., Weber, K. and Wenger, B. (2010) 'Investigating the Accessibility of State Government Web Sites in Maryland', in Langdon, P. M., Clarkson, P. J. and Robinson, P., eds., *Designing Inclusive Interactions - Proceedings of the 2010 Cambridge Workshop on Universal Access and Assistive Technololgy*, Springer London, 69-78.

Lazar, J., Feng, J., Brooks, T., Melamed, G., Wentz, B., Holman, J., Olalere, A. and Ekedebe, N. (2012) 'The SoundsRight CAPTCHA: an improved approach to audio human interaction proofs for blind users', in *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, Austin, Texas, USA, 2208385: ACM, 2267-2276.

Leporini, B. and Paternò, F. (2004) 'Increasing usability when interacting through screen readers', *Universal Access in the Information Society,* 3(1), 57-70.

Leporini, B. and 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.

Leuthold, S., Bargas-Avila, J. A. and Opwis, K. (2008) 'Beyond web content accessibility guidelines: Design of enhanced text user interfaces for blind internet users', *Int. J. Hum.-Comput. Stud.,* 66, 257-270.

Loiacono, E. and McCoy, S. (2004) 'Web site accessibility: an online sector analysis', *Information Technology & People,* 17(1), 87 - 101.

Lopes, R. and Carriço, L. (2010) 'Macroscopic characterisations of Web accessibility', *New Review of Hypermedia and Multimedia,* 16(3), 221 - 243.

Lopes, R., Gomes, D. and Carriço, L. (2010) 'Web not for all: a large scale study of web accessibility', in *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*, Raleigh, North Carolina, ACM, 1-4.

Mankoff, J., Fait, H. and 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*, Portland, Oregon, USA, 1054979: ACM, 41-50.

McCarthy, J. and Swierenga, S. (2010) 'What we know about dyslexia and Web accessibility: a research review', *Universal Access in the Information Society,* 9(2), 147-152.

Mohamad, Y., Stegemann, D., Koch, J. and Velasco, C. A. (2004) 'imergo: Supporting Accessibility and Web Standards to Meet the Needs of the Industry via Process-Oriented Software Tools' in Miesenberger, K., Klaus, J., Zagler, W. and Burger, D., eds., *Computers Helping People with Special Needs*, Springer Berlin / Heidelberg, 628-628.

Molich, R. (2007) *Usable Web Design,* Narayana Press.

Monk, A. (1993) *Improving your human-computer interface : a practical technique, BCS practitioner series,* Hemel Hempstead, Herts.: Prentice Hall International.

Nielsen, J. (2000) *Designing web usability: the practice of simplicity,* Indianapolis: New Riders.

Norman, D. (1988) *The design of everyday things,* 1st ed., New York: Basic Books.

Office for Disability Issues (2011) 'Disability facts and figures', [online], available: http://odi.dwp.gov.uk/disability-statistics-and-research/disability-facts-and-figures.php [accessed 20/07/2012].

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.

Pederick, C. (2011) 'Web Developer Firefox Toolbar', [online], available: https://addons.mozilla.org/en-US/firefox/addon/web-developer/ [accessed 06/01/2012].

Petrie, H. and Bevan, N. (2009) 'The evaluation of accessibility, usability and user experience' in Stephanidis, C., ed. *The Universal Access Handbook*, CRC Press, 20-I - 20-XXX.

Petrie, H., Hamilton, F. and King, N. (2004) 'Tension, what tension?: Website accessibility and visual design', in *Proceedings of the 2004 international cross-disciplinary workshop on Web accessibility (W4A)*, New York, NY, USA, ACM, 13-18.

Petrie, H., Hamilton, F., King, N. and Pavan, P. (2006) 'Remote usability evaluations With disabled people', in *Proceedings of the SIGCHI conference on Human Factors in computing systems*, Montréal, Québec, Canada, 1124942: ACM, 1133-1141.

Petrie, H. and Kheir, O. (2007) 'The relationship between accessibility and usability of websites', in *Proceedings of the SIGCHI conference on Human factors in computing systems*, New York, NY, USA, ACM, 397-406.

Petrie, H. and 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 2012 ACM annual conference on Human Factors in Computing Systems*, Austin, Texas, USA, 2208363: ACM, 2107-2116.

Petrie, H., Power, C., Swallow, D., Velasco, C. A., Gallagher, B., Magennis, M., Murphy, E., Collin, S. and Down, K. (2011) 'The Value Chain for Web Accessibility: Challenges and Opportunities', in Weber, G., Petrie, H. and Darzentas, J., eds., *Accessible Design in the Digital World 2011*, Lisbon, Portugal, CEUR Workshop Proceedings, 2-12.

Potter, A. (2002) 'Accessibility of Alabama government Web sites', *Journal of Government Information,* 29(5), 303-317.

Power, C., Petrie, H., Freire, A. P. and Swallow, D. (2011) 'Remote evaluation of WCAG 2.0 techniques by web users with visual disabilities', in *Proceedings of the 6th international conference on Universal access in human-computer interaction: design for all and eInclusion - Volume Part I*, Orlando, FL, 2022625: Springer-Verlag, 285-294.

Power, C., Petrie, H. and Swallow, D. (2009) 'Functional user requirements for accessibility in education', *EU4ALL Project - Report D 1.3.6* [online], available: http://www.eu4all-project.eu/sites/default/files/content-files/page/11/03/d136-functionalspecsuserrequirement-final.pdf [accessed 09/02/2012].

Rello, L., Kanvinde, G. and Baeza-Yates, R. (2012) 'Layout guidelines for web text and a web service to improve accessibility for dyslexics', in *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility*, Lyon, France, 2207048: ACM, 1-9.

Ridpath, C. and Chisholm, W. (2000) 'Techniques For Accessibility Evaluation And Repair Tools', *Web Accessibility Initiative (WAI), World Wide Web Consortium (W3C)* [online], available: http://www.w3.org/TR/AERT [accessed 09/11/2011].

Ruth-Janneck, D. (2011a) 'An integrative accessibility engineering approach using multidimensional classifications of barriers in the web', in *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility*, Hyderabad, Andhra Pradesh, India, ACM, 1-7.

Ruth-Janneck, D. (2011b) 'Experienced Barriers in Web Applications and Their Comparison to the WCAG Guidelines', in Weber, G., Petrie, H. and Darzentas, J., eds., *Workshop on Accessible Design in the Digital World 2011*, Lisbon, Portugal, CEUR Workshop Proceedings, 13-29.

Rømen, D. and Svanæs, D. (2008) *Evaluating web site accessibility: validating the WAI guidelines through usability testing with disabled users,* translated by New York, NY, USA: ACM, 535-538.

Rømen, D. and Svanæs, D. (2011) 'Validating WCAG versions 1.0 and 2.0 through usability testing with disabled users', *Universal Access in the Information Society,* (Online first), 1-11.

Santana, V. F., Oliveira, R., Almeida, L. D. A. and Baranauskas, M. C. C. (2012) 'Web accessibility and people with dyslexia: a survey on techniques and guidelines', in *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility*, Lyon, France, 2207047: ACM, 1-9.

Seeman, L. (2006) 'Formal Objection to WCAG 2.0', [online], available: http://lists.w3.org/Archives/Public/w3c-wai-gl/2006AprJun/0368.html [accessed

Shneiderman, B. (2000) 'Universal usability', *Commun. ACM,* 43, 84-91.

Shneiderman, B. (2003) *Promoting universal usability with multi-layer interface design,* translated by New York, NY, USA: ACM, 1-8.

Sloan, D. (2006) *The Effectiveness of the Web Accessibility Audit as a Motivational and Educational Tool in Inclusive Web Design*, unpublished thesis (PhD Thesis), University of Dundee.

Sloan, D., Gregor, P., Booth, P. and Gibson, L. (2002) 'Auditing accessibility of UK Higher Education web sites', *Interacting with Computers,* 14(4), 313-325.

Sloan, D., Heath, A., Hamilton, F., Kelly, B., Petrie, H. and Phipps, L. (2006) 'Contextual web accessibility - maximizing the benefit of accessibility guidelines', in *Proceedings of the 2006 international cross-disciplinary workshop on Web accessibility (W4A): Building the mobile web: rediscovering accessibility?*, Edinburgh, U.K., 1133242: ACM, 121-131.

Stanton, N. (2005) *Human factors methods : a practical guide for engineering and design,* Aldershot: Ashgate Pub. Co.

Thatcher, J., Waddell C.D. Henry S.L. Swierenga S. Urban M.D. Burks M. Regan B. and Bohman, P. (2003) *Constructing accessible web sites,* San Francisco: glasshaus.

Theofanos, M. F. and Redish, J. (2003) 'Bridging the gap: between accessibility and usability', *interactions,* 10(6), 36-51.

Theofanos, M. F. and Redish, J. (2005) 'Helping Low-vision and Other Users with Web Sites That Meet Their Needs: Is One Site for All Feasible?', *Technical Communication,* 52(1), 9-20.

Treviranus, J., McCathieNevile, C., Jacobs, I. and Richards, J. (2000) 'Authoring Tool Accessibility Guidelines 1.0 ', [online], available: http://www.w3.org/TR/ATAG10/ [accessed 23/01/2012].

U.S. Department of Health & Human Services (2006) 'The Research-Based Web Design & Usability Guidelines', [online], available: http://www.usability.gov/guidelines/ [accessed 20/07/2012].

UK Government (2010) 'Equality Act 2010', [online], available: http://www.legislation.gov.uk/ukpga/2010/15/contents [accessed 10/08/2012].

US Government (2011) 'US Rehabilitation Act - Section 508', [online], available: http://www.section508.gov/ [accessed 28/11/2011].

Vanderheiden, G. (2009) 'Quantification of Accessibility: Guidance for More Objective Access Guidelines

Universal Access in Human-Computer Interaction. Addressing Diversity' in Stephanidis, C., ed., Springer Berlin / Heidelberg, 636-643.

Velleman, E., Koch, J., Velasco, C. A., Strobbe, C., Nietzio, A. and Snaprud, M. (2007) 'Unified Web Evaluation Methodology (UWEM 1.2 Tests)', [online], available: http://www.wabcluster.org/uwem1_2/UWEM_1_2_TESTS.pdf [accessed 06/01/2012].

Watanabe, T. (2007) 'Experimental evaluation of usability and accessibility of heading elements', in *Proceedings of the 2007 international cross-disciplinary conference on Web accessibility (W4A)*, Banff, Canada, 1243473: ACM, 157-164.

Web Accessibility Benchmarking Cluster (2007) 'Unified Web Evaluation Method 1.2', [online], available: http://www.wabcluster.org/uwem1_2/ [accessed 28 January 2010].

Web Accessibility in Mind (2009a) 'Screen Reader User Survey #2 Results', [online], available: http://webaim.org/projects/screenreadersurvey2/ [accessed 22/11/2011].

Web Accessibility in Mind (2009b) 'Survey of Preferences of Screen Readers Users', [online], available: http://webaim.org/projects/screenreadersurvey/ [accessed 22/11/2011].

Web Accessibility in Mind (2011) 'Screen Reader User Survey #3 Results', [online], available: http://webaim.org/blog/screen-reader-user-survey-3-results/ [accessed 22/11/2011].

Web Accessibility Initiative (2008) 'Web Content Accessibility Guidelines 2.0 – Implementation report - results', [online], available: http://www.w3.org/WAI/GL/WCAG20/implementation-report/results [accessed 09/11/2011].

Web Accessibility Initiative (2012) 'PDF Techniques for WCAG 2.0 ', [online], available: http://www.w3.org/TR/WCAG-TECHS/pdf.html [accessed 17/01/2012].

World Wide Web Consortium (2012) 'Markup validation service', [online], available: http://validator.w3.org/ [accessed 06/01/2012].

Yesilada, Y., Brajnik, G. and Harper, S. (2009) *How much does expertise matter?: a barrier walkthrough study with experts and non-experts,* translated by New York, NY, USA: ACM, 203-210.

Yesilada, Y., Brajnik, G., Vigo, M. and Harper, S. (2012) 'Understanding web accessibility and its drivers', in *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility*, Lyon, France, 2207027: ACM, 1-9.

Yuquan, S. (2007) 'The accessibility of Chinese local government Web sites: An exploratory study', *Government Information Quarterly,* 24(2), 377-403.

Zarach, V. (2002) 'Ten Guidelines for Improving Accessibility for People with Dyslexia. ', *CETIS University of Wales Bangor* [online], available: http://wiki.cetis.ac.uk/Ten_Guidelines_for_Improving_Accessibility_for_People_with_Dyslexia [accessed 29/06/2011].