Search and find?
An accessibility study of dyslexia and information retrieval

by

Gerd Berget

Thesis submitted for the degree

Of Philosophiae Doctor

Faculty of Mathematics and Natural Sciences

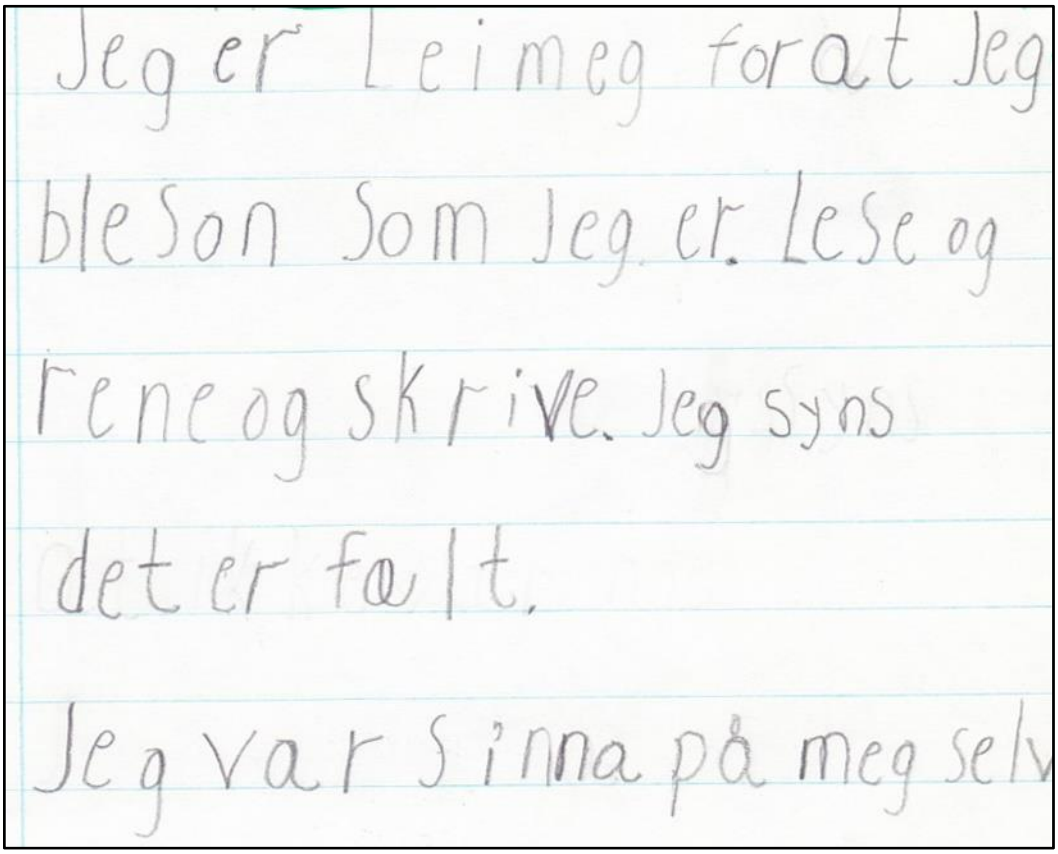
University of Oslo

2016

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 1762

ISSN 1501-7710

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.



Jeg er Leimeg forat Jeg
bleSon Som Jeg er. Lese og
rene og skrive. Jeg syns
det er fælt.
Jeg var Sinna på meg selv

«*I am sorry I turned out the way I am. Read and calculate and write. I think it is awful. I was mad at myself*»
Girl, 12 years old (the text contains several misspellings)

Abstract

Dyslexia is prevalent in 5-10% of the population. Although studies have addressed several issues related to Web accessibility and dyslexia, little attention has been directed at the interaction with search user interfaces. Consequently, there is a gap in the knowledge of how information retrieval is influenced by reading and spelling difficulties.

This study set out to investigate how dyslexia affects information search. A total of 42 students (21 dyslexics and 21 controls) completed three experiments; one visual search experiment and two information retrieval experiments in the Web search engine Google and the academic library catalogue Bibsys Ask. Eye tracking and screen recording documented the searching.

The main purpose was to explore the impact of dyslexia on information searching. Further, it was an aim to investigate whether more accessible design may counteract potential negative effects of dyslexia on information retrieval without reducing the accessibility for other users. Searching has become a common activity for most online users. Creating more accessible search user interfaces may thus be an important measure to achieve both a more universally designed Web in general, and also enhance the accessibility of libraries in particular. An overall aim of this study was to contribute with empirical knowledge that may support the design of more accessible search systems for both dyslexic and non-dyslexic users.

Dyslexia was found to have a significant effect on query-formulation. However, a high tolerance for errors seems to better counteract the effect of dyslexia than the presence of query-building aids. Actually, the search system with a high fault tolerance removed the negative effect of dyslexia on search performance, which indicates that all interfaces that includes search facilities may be equally accessible for dyslexics as non-dyslexics if the interface is adequately designed.

Help functions during query-formulation were not applied significantly more by dyslexics than non-dyslexics due to, among others, an intense focus on the keyboard during query input. Consequently, there may be a need for a different search user interface design, better keyboard skills or a different keyboard design if this function should be useful for dyslexics. Moreover, a language option regarding the query suggestions in the autocomplete function may be expedient. It was also found that autocomplete was mostly utilised for dictionary purposes, and not to enhance the query formulation in systems with a high tolerance for spelling errors.

This study also investigated whether graphic content in result lists may be useful for dyslexic users. The main conclusion regarding result list interfaces was that the usefulness of graphic content relates to, among others, layout and spacing in terms of whether the two modalities are presented within the central visual field concurrently. It is not possible based on the findings to fully support guidelines suggesting to add images or icons to enhance the accessibility for dyslexics, since the usefulness seem to depend on presentation and layout. However, when dual-modalities are presented in a list-layout distanced so that icons and words are not in the central visual field concurrently, graphic content may seem helpful and also seems to counteract the negative impact of dyslexia on visual search tasks.

The findings have implications for Web accessibility guidelines applied to search user interfaces, and may contribute to more accessible search. Both dyslexic and non-dyslexic users shared similar search strategies and encountered some of the same difficulties in all the three experiments. Consequently, measures that enhance the accessibility for dyslexics may also benefit non-dyslexic users.

Acknowledgements

This project has been financially supported by the Norwegian ExtraFoundation for Health and Rehabilitation through EXTRA funds grant 2011/12/025.

I wish to thank my main supervisor Frode Eika Sandnes. Frode has believed in me and the project from the start, been a great motivator, invested a lot of time and asked an approximately large amount of critical questions on the way. Your effort, friendship and our scholarly debates have been highly appreciated. I also want to thank Jo Herstad, who has a wonderful view on life and has contributed with valuable feedback, interesting philosophical discussions in addition to showing impressive one-wheel bicycling and juggling skills. Fiona Mulvey has introduced me to the amazing world of eye tracking and been an excellent hostess during my visits at the eye tracking lab in Lund. Norun Christine Sanderson has helped with piloting, provided valuable input and moral support.

I wish to thank Dyslexia Norway for helping with recruitment and administration related to funding, especially Åsne Midtbø Aas, Kine Hvinden Dahl and Caroline Solem. I am also grateful to the participating students for meaningful discussions and pleasant meetings at the lab. Gro Horgen Vikesdal has been a good advisor for visual acuity tests.

I also wish to thank colleagues at the Institute of information technology and Department of Archivistics, Library and Information Science, especially Siri Fagernes, Hugo Lewi Hammer, Laurence Habib and Nils Pharo.

A special thanks to Michel Habib for introducing me to the world of dyslexics. Michel provided me with valuable knowledge and motivation before starting this project. I am also grateful to the anonymous reviewers of the papers for feedback, and Alma Leora Culén and Petter Nilsen for contributing with valuable comments on my trial defence.

I wish to thank Hanne Hognestad and our «beaver scouts» in Hyggen for a much needed break in the forest every Tuesday afternoon. Last but not least, I want to thank my wonderful family; my father Per Magne for teaching me that a focus on details is always important and for being the person I can call and ask about «everything», my mom Ingrid for lots of motivating phone calls, support and interesting discussions.

I also want to thank my amazing grandmother Ellen Johanne, my always supportive brother Stig, his wonderful wife Christina and my beautiful nieces Sofie and Pernille for believing in me. I also want to thank my in-laws, Ann-Mari and Jan for great support. And of course, my kind and patient boyfriend Sten Are, for asking questions which have made me see things in new perspectives and for bearing with me when my PhD project might have been mentioned a bit too many times during family dinners. Finally, I want to thank my son Østen, the light of my life, who is always there with a hug and a smile.

Preface

This thesis is submitted for the PhD degree at Department of informatics, Faculty of Mathematics and Natural Sciences, University of Oslo, Norway. Experiments have been conducted at the Institute of Information Technology, Faculty of Technology, Art and Design at Oslo and Akershus University College of Applied Sciences, Norway. In addition, a part of the piloting was carried out at Department of Archivistics, Library and Information Science, Faculty of Social Sciences at Oslo and Akershus University College of Applied Sciences, Norway. Frode Eika Sandnes[a] has been the main supervisor. Jo Herstad[b], Fiona Mulvey[c] and Norun Christine Sanderson[a] were co-supervisors.

This thesis consists of an introduction and four papers (A-D) reporting the research findings. In addition, two related papers (E-F) are included in the appendices:

A. Berget, G. & Sandnes, F. E. (2015). The effect of dyslexia on searching visual and textual content: Are icons really useful?. In M. Antona & C. Stephanidis (Eds.), *Universal Access in Human-Computer Interaction: Access to Learning, Health and Well-being: 9th International Conference, UAHCI 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings, Part III* (pp. 616-625). Cham: Springer.
B. Berget, G., Mulvey, F. & Sandnes, F. E. (2016). Is visual content in textual search interfaces beneficial to dyslexic users?. *International Journal of Human-Computer Studies, 92-93, 17-29*.
C. Berget, G. & Sandnes, F. E. (2015). Searching databases without query-building-aids: Implications for dyslexic users. *Information Research, 20*(4), paper 689.
D. Berget, G. & Sandnes, F. E. (Accepted 2015). Do autocomplete functions reduce the effect of dyslexia on information searching behavior? The case of Google. *Journal of the American Society for Information Science & Technology*. Accepted 2015, early-view published online October 2015.
E. Berget, G. & Sandnes, F. E. (2015). On the understandability of public domain icons: Effects of gender and age. In M. Antona & C. Stephanidis (Eds.), *Universal Access in Human-Computer Interaction: Access to Today's Technologies, 9th International Conference, UAHCI 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings, Part I* (pp. 387-396). Cham: Springer.
F. Berget, G., Herstad, J. & Sandnes, F. E. (2016). Search, read and write: An inquiry into Web accessibility for dyslexics. In H. Petrie, J. Darzentas, T. Walsh, D. Swallow, L. Sandoval, A. Lewis & C. Power (Eds), *Universal Design 2016: Learning from the Past, Designing for the Future* (pp. 450-460). Amsterdam: IOS Press.

Finally, Appendix G contains an approval letter from the Norwegian Social Science Data Services (NSD) confirming the project's compliance with ethical standards.

[a]Institute of Information Technology, Faculty of Technology, Art and Design, Oslo and Akershus University College of Applied Sciences, Oslo, Norway
[b]Research Group for Design of Information Systems, Faculty of Mathematics and Natural Sciences, University of Oslo, Norway
[c]Humanities Lab, The Faculty of Humanities and Theology, Lund University, Sweden

Contents:

Chapter 1: Introduction

Online information searching has become a common activity, and much content on the Web is retrieved through search user interfaces, library catalogues or academic databases. However, searching was available before the Web was introduced in the middle of the 1990s. In the context of libraries, the first online public access catalogues were introduced as an alternative to the traditional card catalogues at the end of the 1980s (Husain & Ansari, 2006). According to Schwartz (1998), the Internet became available for scholars and others communities towards the end of the 1980s, when Telnet was applied to search online directories and libraries.

Bates (1989) suggested in 1989 that "*Soon there will be something approaching whole libraries accessible by computer*". Today most libraries offer their patrons a variety of digital services, including searching the library catalogue and access to a variety of online books and journals. Librarians, especially in academic libraries, usually offer courses and personal help in searching. However, several libraries have expanded their services by offering self-service facilities after regular opening hours for their users and are therefore not always available when people are searching for information. Consequently, library users, especially in academic libraries, are expected to locate and retrieve relevant literature by themselves.

Search has always been an important part of the Web. The hypertext paradigm, where documents are connected through hyperlinks, is a basic foundation for the Web. However, when Berners-Lee and Cailliau (1992) initially discussed the World Wide Web project, they did not consider hyperlinks sufficient to navigate large amounts of Web content. Consequently, Berners-Lee and Cailliau (1992) suggested that searchable indexes should be added to the hypertext model. The main idea was that users could provide keywords or other search criteria, and be presented with a document containing hyperlinks to relevant documents.

Two decades later the Web is believed to comprise several billion Web pages, depending on the estimation algorithm applied (van der Bosch, Bogers & de Kunder, 2016). Browsing and searching are still regarded as fundamental means of interaction on the Web (Dörk, Williamson & Carpendale, 2012). According to Cutrell and Guan (2007) users have become highly dependent on search engines to locate information, and Gossen, Hempel and Nürnberger (2013) report that online searching has become a part of most people's daily routines, both among children and adults. For instance, in 2012, 73% of Americans reported a daily use of search engines (Purcell, Brenner & Rainie, 2012), and billions of Web searches are conducted worldwide every day (Internet Live Stats, n.d.).

The Web has a diverse content, from personal blogs to peer-reviewed academic papers. Jamali and Hasadi (2010) reported that online information is ordinarily retrieved through general search engines, library catalogues and academic article databases. Association of College and Research Libraries (2000) has emphasised information literacy as a basis for lifelong learning, which includes abilities such as to locate, evaluate and efficiently use information. University course reading lists typically include a considerable amount of online articles, and lecturers expect students to locate reliable information sources as background material for assignments, such as a bachelor or master theses. Consequently, a frequent use of search systems is necessary to locate academic literature and to complete a degree.

Information searching may be difficult. According to Aula, Khan and Guan (2010) many people struggle occasionally with online information retrieval. Taylor (1968) was one of the first to address the difficulties with transforming vague information needs into concrete queries. It has also been suggested that insufficient information searching skills or inaccessible search user interfaces may cause difficulties for end users (Hearst, 2009). For instance, layouts can be confusing, or the system may require the user to formulate advanced queries with technical language or advanced search techniques.

Certain search difficulties may be resolved through formal training and extensive search experience. Nevertheless, challenges caused by inadequate user interface design should rather be resolved by altering the actual systems. This is in accordance with the user-centred paradigm introduced by Norman, where the aim is to develop and test products around user behaviour, and not demand that users change their behaviour to adapt to the products (Abras, Maloney-Krichmar & Preece, 2004).

Search user interfaces intended for the general public should be understandable for users of all cultures, ages and backgrounds and apply to different information needs and contexts. Hearst (2009) describes the basic search option in most user interfaces as a relatively simple design, often merely consisting of a text field for query input and a submit button. However, end users' motoric, sensory and cognitive characteristics vary. It is therefore claimed that functional levels should be taken into account when considering the usability of a search system (Disability Rights Commission, 2004).

The term Web accessibility is commonly applied when the Web is discussed in a context of disabilities (Petrie & Kheir, 2007). Web accessibility refers to the extent in which a Web site is usable for users with impairments. More specific, The Web Accessibility Initiative defines Web Accessibility to entail that «*people with disabilities can perceive, understand, navigate, and interact with the Web, and that they can contribute to the Web*» (W3C, 2005).

The importance of designing an accessible Web was addressed by Berners-Lee, the founder of the Web. Berners-Lee stated, «T*he power of the Web is in its universality. Access by everyone regardless of disability is an essential aspect*» (W3C, 2014). Access to information as a fundamental right for all people is also embedded by UN in the Convention on the Rights of Persons with Disabilities (UN, 2006). Story, Mueller and Mace (1998) have claimed that an understanding of human diversity is critical to accomplish accessibility for all. Such an understanding includes knowledge about how user characteristics such as cognition, vision, hearing, speech, body function and mobility affect the use of technology.

In the context of cognition, it has been suggested that dyslexics struggle with Web navigation (Al-Wabil, Zaphiris & Wilson, 2007). It is also suggested that dyslexics have difficulties with information searching (Habib et al., 2012), and that there is a need to further investigate how dyslexia affects information search (Hepworth, 2007; MacFarlane et al, 2010; MacFarlane, Albrair, Marshall & Buchanan, 2012). According to the World Wide Web Consortium, the existing guidelines for Web accessibility (WCAG) do not fully accommodate users with reading impairments (W3C, 2008). Consequently, there may also be a need for the revision of existing accessibility guidelines to also accommodate dyslexic users.

The objective of this thesis is to acquire empirical data on how dyslexia affects information search. The interaction between users and search systems is studied with an aim to uncover potential shortcomings in existing search user interfaces. These findings can be used in the development of Web accessibility guidelines to better accommodate dyslexic users. McCarthy and Swierenga (2010) claimed that measures which increase the accessibility for dyslexics also benefit other users, for instance by presenting clear and readable text or more navigable Web sites. Consequently, dyslexics may also be an applicable user group to study how to improve the interaction between users and search systems in general.

Section 1.1 addresses Web accessibility, including a short overview of the disability discourse and terminology. Section 1.2 introduces the main research question and three hypotheses that constitute the basis for the study. This is followed by Section 1.3 that describes the theory, Section 1.4 addresses methodology, and Section 1.5 includes a short description of the experiments. In the final section, 1.6, the overall structure of the thesis is outlined.

1.1 Web accessibility

Users have received increased attention in the research and development of computer systems over the last decades. Human computer interaction (HCI) has developed into a broad research field, and according to Abras et al. (2004), the user-centred design tradition is now a highly accepted approach. In user-centred design, user requirements are incorporated in the software developmental process.

Another much applied paradigm in HCI is the Scandinavian tradition, also referred to as participatory design, where users are even more actively participating in the design process (Fuchs & Obrist, 2010). User involvement is regarded to have advantages such as a better understanding of the users and their contexts and a reasonable allocation of tasks between the system and user in an iterative design process, where feedback from users is incorporated during the development process (Maguire, 2001).

User diversity and accessibility have caught attention among human-computer interaction researchers and practitioners, and a shift towards universal design was reported in the literature more than a decade ago (Newell & Gregor, 2000). The Center for Universal Design (2008) applies the following definition of universal design: «*Universal design is the design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design*».

Henry, Abou-Zahra and Brewer (2014) claim that universally designed user interfaces has become a goal in much user-centred design. Universal design and accessibility are also included in national legislations and international human rights treaties such as the Convention on the Rights of Persons with Disabilities (UN, 2006). The UN convention states among others that States Parties should «*promote universal design in the development of standards and guidelines*» and «*provide accessible information to persons with disabilities*».

Attempts have been made to promote universal design on the Web, for instance the development of accessibility guidelines. Such guidelines typically originate from theory-driven guidelines derived from academia, and are based on practical experiences in the industry (Zaphiris, Kurniawan & Ghiawadwala, 2007). The Web Content Accessibility Guidelines, also referred to as WCAG (W3C, 2008), has now become the de facto standard for many Web designers and developers. According to W3C (2008), WCAG should assist developers in implementing universal design on the Web and be a shared standard of minimum requirements which must be fulfilled to claim that a product is universally designed.

Accessibility guidelines are developed with the purpose of accommodating a wide range of users. In the introduction to WCAG, W3C (2008) lists blindness and low vision, deafness and hearing loss, learning disabilities, cognitive limitations, limited movement, speech disabilities, photosensitivity and a combination of these impairments. The objectives behind these guidelines are, among others, barrier free and equal access for all. However, Newell and Gregor (2000) emphasise that diverse user groups may have diametrically opposite needs. For instance, while deaf users depend on visual content, visually impaired users rely on auditory content. In contrast, some preferences are shared by several user groups. For instance, voice captions for text may accommodate both visually impaired users and people with reading disabilities, while written information may be preferred both by deaf users and people in noisy environments.

Universal design incorporates all users. In contrast, Story et al. (1998) argue that people with disabilities were previously not sufficiently incorporated in attempts to develop a more accessible Web. It was not uncommon with special design features or specially designed Web pages for impaired users, thus excluding people with disabilities from the main design. Such pages often had a different design than the original Web site, and it entailed increased costs to maintain two parallel user interfaces (Story et al., 1998).

Universal design emerged as an alternative strategy focusing on the development of one accessible interface, which included as many people as possible, despite physical, cognitive and situational differences (Henry et al., 2014). According to Newell and Gregor (2000), universal design may be regarded a subfield within the user-centred approach, but has some distinctive characteristics, such as a greater diversity of users.

The main purpose of WCAG (W3C, 2008) is the design of an accessible Web for all. However, user studies have revealed that conformity with these guidelines may not guarantee accessible Web pages for all users (Brajnik, 2009; Rømen & Svanæs, 2012). Shortcomings in WCAG are also addressed in its introduction, where W3C (2008) states that Web pages «*will not be accessible to individuals with all types, degrees, or combinations of disability, particularly in the cognitive language and learning area*». This issue is also addressed in the research literature.

Previous studies have suggested that the Web in general and search systems in particular were inaccessible to many dyslexics (Al-Wabil et al., 2007; Habib et al., 2012; Santana, Oliveira, Almeida & Baranauskas, 2012). Consequently, there may be a need for revised guidelines which also accommodate users with cognitive disabilities, such as dyslexia. Dyslexia is a common reading and writing impairment which affects at least 5-10% of any population (Ahissar, 2007). Consequently, inaccessible Web pages may cause difficulties for many users and is an issue which needs to be addressed.

Although general Web accessibility guidelines exist, alternative Web design guidelines directly targeted at dyslexic users have been discussed by the British Dyslexia Association (n.d.), Zarach (2002) and Santana et al. (2012). Such guidelines promote for instance clear navigation, a broader use of graphics, a clear writing style and a purposeful use of typefaces and font sizes. These guidelines are specifically directed at Web page design, but may also apply to search systems, such as the use of typefaces, applying clear instructions and presenting results in a layout which does not include links in the middle of text.

Guidelines targeted directly at one user group are reported to improve the usability for the specific users addressed (Story, 1998). However, such practise may not be regarded as coherent with the universal design approach, where all user groups should be included in one joint set of guidelines. Further, it would be time-consuming and impractical for designers and developers to consult separate sets of guidelines for each user group.

McCarthy and Swierenga (2010) claim that dyslexics often have not been included in studies on Web accessibility. The lack of consideration for such a large user group has been the main motivation behind this study. Further, dyslexics are reported to be highly suitable test-users. For instance, Dixon (2007) found in a software evaluation study that dyslexic students uncovered a larger number of user interface problems than non-dyslexics. Further, it is argued that symptoms of dyslexia may be experienced by non-dyslexics in certain contexts. For instance, fatigue may negatively affect memory capacity and concentration, and stressful situations may affect spelling abilities (McCarthy & Swierenga, 2010). There is also an overlap in guidelines for dyslexia and other cognitive or physical disabilities. According to McCarthy and Swierenga (2010), dyslexic users constitute an important group for overall Web accessibility.

1.1.1 Novel contribution

Tops, Callens, Lammertyn, Hees and Brysbaert (2012) have reported an increasing number of dyslexic students entering higher education. Consequently, a large number of dyslexics frequently retrieve scholarly information from the Web. However, it is suggested that dyslexics find information retrieval particularly difficult (Habib et al., 2012), and that dyslexia affects several cognitive characteristics, such as working memory, required for searching (MacFarlane et al., 2012).

Several researchers have reported a gap in the dyslexia research regarding information search (Hepworth, 2007; MacFarlane et al., 2010; MacFarlane et al., 2012). Maguire (2001) claims that sufficient user knowledge is an important characteristic of user-centred design. Consequently, it may be difficult to develop suitable search user interface guidelines without proper apprehension of the users. Thus, the novel contribution of this thesis is empirical knowledge on how dyslexia affects information retrieval, with a particular focus on query formulation, result list assessment and fault tolerance.

The findings from this study can be used to support designers in making more accessible search systems. The results may also be relevant background in discussions on how accessibility guidelines can better accommodate dyslexic users, which is an important measure towards universally designed Web sites and libraries. Moreover, a review of WCAG (W3C, 2008), the existing de facto standard for Web Accessibility, is provided in a related paper by Berget, Herstad and Sandnes (2016), see Appendix F.

1.1.2 Dyslexia and accessible search systems

Figure 1 displays a typical cycle of steps to conduct a search. The process is initiated by a desire to acquire information on a topic to satisfy an information need. Scholars debate why and how information needs arise, but according to Wilson (1999) there is a common understanding that an information need is a state of uncertainty that often initiates information search, and is resolved when the requested information is located. Examples of such information needs may be simple questions, such as finding out which album Bob Dylan released in 1969 to more complex issues such as locating all relevant research on a certain cancer treatment.

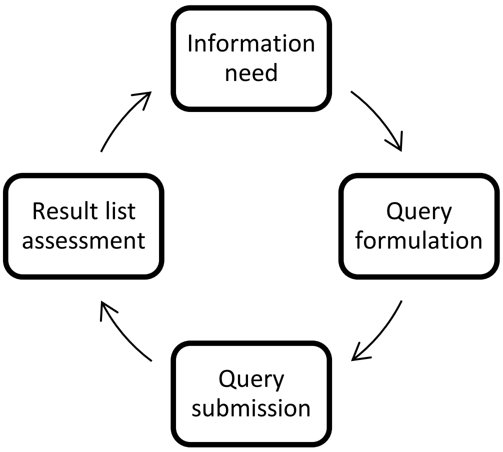


Figure 1: Typical cycle of steps to conduct a search

The actual search process may vary with the complexity of the information need, but there are certain basic stages (see Figure 1) which are frequently incorporated in online information searching models (Case, 2007). When an information need arises, this need is translated into a query, which is inputted and submitted to a search system. A result list containing a set of documents is then returned to the user, who assesses the results according to some criteria for relevance. The search process ends when the user is satisfied, or continues with a new or revised information need until the user is content, or decides to abort the searching.

In a study by Habib et al. (2012) dyslexics mentioned three specific difficulties when searching: formulating useful queries, spelling query terms correctly and assessing large result lists. Gwizdka (2010) reported query formulation to be the phase which generally requires the most cognitive load for all users. In this study it is presumed that formulating queries may be particularly demanding for dyslexics, who are reported to frequently struggle with spelling (Snowling, 2000).

Dyslexia may also cause reduced rapid naming skills, which refers to remembering the names of for instance objects, colours and numbers (Snowling, 2000; Lervåg & Hulme, 2009). In this study it is assumed that difficulties remembering words may cause challenges for dyslexics when formulating queries. Assessing documents may also be demanding. This activity requires reading skills when browsing through large amounts of text. Further, a good short-term memory capacity may be necessary to differentiate between the assessed documents. Both of these cognitive abilities are also often reported to be impaired in dyslexics (Snowling, 2000; Smith-Spark & Fisk, 2007).

Dyslexics have consistently been reported to strive when using physical libraries (Riddick, Farmer & Sterling, 1997; Helland, 2002), and rely on extra library support (Mortimore & Crozier, 2006). However, according to Helland (2002) many dyslexics avoid using the library due to a fear of failure and embarrassment related to using the library services.

The digital library has evolved as a consequence of more electronic publishing and digitalised material. According to Buckland (2008), electronic library services have for many students become a substitution for face-to-face services. Although the digital library has advantages such as round-the-clock services and access to digital content away from the physical library, these changes may also entail that more users must retrieve online information unassisted.

Librarians may act as intermediaries between the user and the search system, and may therefore assist users who struggle with information searching. However, when students are searching scholarly databases away from the library, they have to locate the desired information unaided. This development may be particularly challenging for dyslexics, who may already struggle when using library facilities. Moreover, the need for accessible search systems is even greater when users need to search for information unassisted. Hepworth (2007) pointed out that more empirical knowledge is needed to better understand how to accommodate dyslexic users when designing search systems. The aim is to contribute with such knowledge.

1.1.3 The disability discourse

In the 1980s there was a significant paradigm shift in the disability discourse from a medical model to a social model (Shakespeare, 2013). In the medical model (also referred to as the individual model), a disability was regarded as a discrepancy within an individual compared to normative values. In contrast, the social model defines disability as created by barriers in the society (Oliver, 1996). The paradigm shift did not only affect the definition of disabilities, but also entailed a change of attitudes towards the responsibilities of the society. In the medical model, a disability was regarded as an individual problem which required personal treatment. In contrast, the social model demanded social action, such as changing the physical environment to increase accessibility (Oliver, 1996).

A third perspective is the Nordic relational model. Some researchers argue that this is an alternative model on disabilities, but according to Shakespeare (2004) it may be regarded as a version of the social model. The Nordic relational model is also referred to as the Scandinavian model or the gap-model (see Figure 2). According to the gap-model, disabilities arise due to a mismatch between expectations in society and a person's abilities. Disabilities may be reduced by changing the environment and strengthening individuals (Sosialdepartementet, 2003).

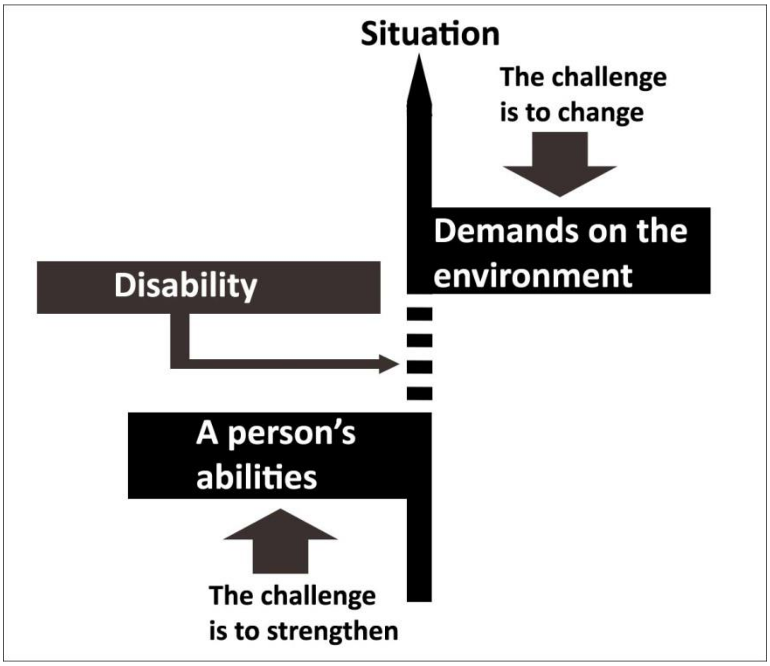


Figure 2: The gap-model (translated reproduction based on Sosialdepartementet, 2003: 9)

Universal design is a philosophy which may comply with both the social model and the gap-model, since it is an attempt to change the society to counteract disabilities. The focus is directed at environmental factors, and the society is expected to design for user diversity. Universal design may be regarded as a counterpart to accessible design. Iwarsson and Ståhl (2003) have argued that accessible design is based on the assumption that the population consists of people with and without disabilities, which must be taken into account when designing products and services. In contrast, universal design is based on the principle that there is only *one* population consisting of individuals with different characteristics and abilities, and design solutions must accommodate the *whole* population.

One reason for using the relational gap-model in this study is the belief that universal design may not be possible to achieve in all contexts for everyone. However, through an increased effort in designing universally designed search user interfaces, search systems may become more accessible to dyslexics. Principles and practises regarding universal design are discussed in more detail in Section 2.1.

1.1.4 Terminology and stigma

The transition from the medical model to the social model did not only affect the disability discourse, but also entailed a change of terminology. According to Oliver (1996), the medical model focused on disability as a counterpart to normality, where non-average bodies were considered abnormal and in many contexts related to the abilities to work. The comparison with normality led to a discourse where users with disabilities were regarded as abnormal, which resulted in stigmatisation. In the social model, physical and social environments are considered as enabling or disabling, and a disability is something that occurs in inadequately constructed societies only. The social model is more politically loaded than the medical model, and the focus has often been directed at marginalization and discrimination. In contrast, the gap-model is reported not to take any position regarding disabilities, but simply states that there will always be a portion of the population who may be regarded as disabled (Sosialdepartementet, 2003).

The paradigm shift in the disability discourse led to a change of terminology. For instance, words referring to *normality* are now regarded as politically incorrect language in the disability discourse, and are no longer used. Another example is the previously applied term *mental retardation*, which is replaced with *intellectual disabilities*. According to Ford, Acosta and Sutcliffe (2013), mental retardation is now regarded as a negatively loaded stigmatizing word. Another example is an ongoing campaign in, among others, social media to replace the term *learning disabilities* with *learning differences* (LDA, 2008). This terminology change emphasises diversity rather than applying one specific learning preference as a norm.

The choice of words when referring to disabilities may be interpreted as positioning in the disability discourse. Consequently, terminology is important. However, selecting the proper terminology may be difficult. In this thesis, there were two contradictory aspects to consider: respectful referrals to dyslexic participants and the readability of the thesis. Concerning dyslexia, Evans (2013) found in a study that nursing students with dyslexia rejected medical language. Of the 12 students interviewed by Evans (2013), seven participants objected to the use of disability and eight objected to the word impairment to describe their dyslexia. The nursing students expressed their attitudes towards dyslexia accordingly: «*I am not a dyslexic person, I'm a person with dyslexia*» (Evans, 2013).

Wordings such as person/participant/user with dyslexia were considered in this thesis. However, such phrasing would have unnecessarily increased sentence lengths and reduced the readability. This was also the feedback from reviewers on the first manuscripts submitted for publication. Consequently, when referring to people with dyslexia, words such as dyslexics, dyslexic users or people/participants with dyslexia are applied. The users without dyslexia are typically referred to as controls, control users, non-dyslexics or users without dyslexia. This seems to be in accordance with the special interest organisation Dyslexia Norway, which among others publish a magazine called *Dyslektikeren* (translates the dyslexic).

The choice of terminology is not intended to be discriminative towards dyslexics, and does not support an attitude that people should be defined solely according to their impairment. However, the purpose of this study is to investigate how dyslexia affects information retrieval and the interaction with search systems. Although dyslexia is just a small part of the dyslexic participants' identity, it is inevitable the main focus in this context. Consequently, the terminology must clearly indicate which of the two participant groups that are being discussed. The presence or absence of dyslexia is therefore applied when referring to the two participant groups with terms such as dyslexics or non-dyslexics.

1.2 Research question

The main topic is how dyslexics interact with search systems. However, information retrieval comprises different activities, such as select proper sources, identify search terms, formulate queries, assess documents and use information (Wilson, 1999). Further, users search at different devices, such as desktop computers, smartphones and tablets, and also in different contexts (Case, 2007). It is not possible to cover all these variables in one study. Consequently, certain limitations are required.

Several researchers have suggested that dyslexia may negatively affect information retrieval (Hepworth, 2007; MacFarlane et al., 2010; Habib et al., 2012; MacFarlane et al., 2012). This study investigates two specific steps of information searching, namely query formulation and result list assessment. The aim was to explore how dyslexia affects information searching behaviour, and the following research question constitutes the basis for the study:

> *How are query-formulation and result list assessment in online search systems affected by dyslexia?*

A total of three main hypotheses were formulated for this research question, one regarding result lists (H_1) and two addressing query-formulation (H_2 and H_3). In addition, a number of more detailed hypotheses were formulated for each experiment. These hypotheses are closely related to H_1, H_2 and H_3, but do more directly address explicit measures such as time usage, number of queries needed to solve a specific task and the number of misspellings for that particular stimulus or search system applied in the experiment. The more detailed hypotheses are presented in Chapter 4 with the research findings, and in the papers attached as Appendices A-D.

1.2.1 H_1 – Result list content

A set of characteristics are reported to be common among dyslexics, such as such as impaired reading and writing skills (Snowling, 2000) and a reduced short-term memory capacity (Smith-Spark & Fisk, 2007). A basis for the research question was that the cognitive profile of dyslexia could affect information retrieval. One assumption was that result list assessment requires, among others, reading skills, since a lot of text must be processed during the evaluation of retrieved documents. Consequently, the first hypothesis addressed whether graphic content can reduce the work load associated with navigating result lists.

According to Prado, Dubois and Valdois (2007), researchers do not seem to agree on the effect of dyslexia on visual search performance. This is supported by Brante, Olander and Nyström (2013), who reported inconclusive findings in previous research regarding the relationship between textual and graphic content. The first hypothesis (H_1) addressed visual search skills in graphic versus textual content and in one-modality versus dual-modality layouts.

Shaywitz and Shaywitz (2005) mention word decoding difficulties as one of the main characteristics associated with dyslexia. It was therefore suggested in the first hypothesis that graphic content could reduce the requirements for reading and support the result list assessment. Santana et al. (2012) reported that dyslexics prefer textual content over graphic content. However, researchers seem to disagree on the usefulness of dual-modalities. This issue is discussed further in Section 2.5.3.

In dual-modality interfaces there is a trade-off between the number of modalities and the set-size. Adding a modality increases the number of objects displayed, which is reported to negatively affect cognitive workload and performance such as response times (Jerde, Ikkai & Curtis, 2011). Moreover, Moores, Cassim and Talcott (2011) found search performance among dyslexics to decrease with larger set-sizes. However, Vidyasagar and Pammer (1999) reported that this set-size effect has not been found to appear until a significant number of items are displayed. The contradicting results reported in the research literature confirm the need for research on search performance among dyslexics in dual-modality displays. Consequently, H_1 was formulated as follows:

> **H_1**: *Graphic content in result lists improves search performance among dyslexic users*

The hypothesis was explored in a visual search experiment through the use of icons. Niemelä and Saarinen (2000) reported a general preference for icons in the general population. Snowling (2000) have emphasised word decoding difficulties among dyslexics. It was therefore assumed that dyslexic users may have a preference for icons over text.

1.2.2 H_2 – Query-building aids

Query-building aids are reported to be a common method for assisting users during query-formulation (Hearst, 2009), such as displaying query suggestions while the user inputs the query (autocomplete), make the user aware of misspellings, correct spelling errors, display results immediately during query input, avoid empty result lists and balance the user control with automated actions. The second hypothesis (H_2) addresses the use of such help functions.

Snowling (2000) claims that dyslexics frequently misspell words. The background for H_2 was an assumption that query-building aids may be particularly supportive for users who struggle with correct spelling, for instance because these functions suggest search terms, thus reducing the need for inputting the words correctly. Further, Catts, Gillispie, Leonard, Kail and Miller (2002) claim that reduced rapid naming skills, which refers to difficulties with remembering the proper names of for instance objects, are also commonly found among dyslexics. It was therefore assumed that difficulties recalling words could lead to a more frequent use of the autocomplete function among dyslexics. Another query-building aid which was assumed to assist dyslexic users was feedback on spelling errors, such as the red underlining of misspelled words in Google. The second hypothesis was defined as follows:

> **H_2**: *Dyslexic users rely more on query-building aids than non-dyslexics*

H_2 addresses the general utilisation of query-building aids with the purpose of revealing which functions are actually applied by dyslexic users.

1.2.3 H_3 – Tolerance for errors

Misspelling is reported to constitute one of the most prominent characteristics of dyslexia (Snowling, 2000). Hearst (2009) has presented general design criteria for search systems. These criteria include a point specifically related to reducing errors, and includes, among others, avoiding empty result lists through spelling correction. However, the tolerance level for misspellings in existing search systems varies. Some search systems have a high tolerance for spelling errors, such as the general search engine Google, while other systems provide no results for misspelled queries. An example of the latter is the library catalogue Bibsys Ask, which is applied in Norwegian academic and research libraries.

The third hypothesis (H_3) was based on the assumption that the tolerance level for spelling errors directly affects search performance among dyslexics:

> ***H_3**: The effect of dyslexia on search performance relates to the system's tolerance for errors*

The purpose of H_3 was to identify whether more tolerable systems may counteract the effect of dyslexia on information retrieval. The answer to H_3 may also have implications concerning system preferences among dyslexic users, since dyslexics may favour search systems which allow for misspellings.

1.3 Theory

The theoretical basis for this study mainly relies on HCI research and library and information science. Within the universal design paradigm, several studies have addressed user knowledge, user diversity and accessibility guidelines. It has also been debated whether universal design is an achievable goal, and what role universal design should have within HCI. A selection of dyslexia research has also been included, for instance discussions on how to define dyslexia, studies of the cognitive profile and research related to Web accessibility and information retrieval. Visual search is an important component of information retrieval, and is included among others in the context of basic eye movements and cognition, but also more specifically regarding dyslexics and their visual search skills.

The main purpose is to investigate the interaction between dyslexics and search user interfaces. Researchers have addressed topics such as how information needs arise, how information needs are resolved and the interaction between users and search systems (Case, 2007). Wilson (1999) introduced the Nested Model of Conceptual Areas (see Figure 3), that aims at visualising the most relevant concepts in information seeking and the relationship between them. This model can be used to define which part of information seeking behaviour addressed in a study.

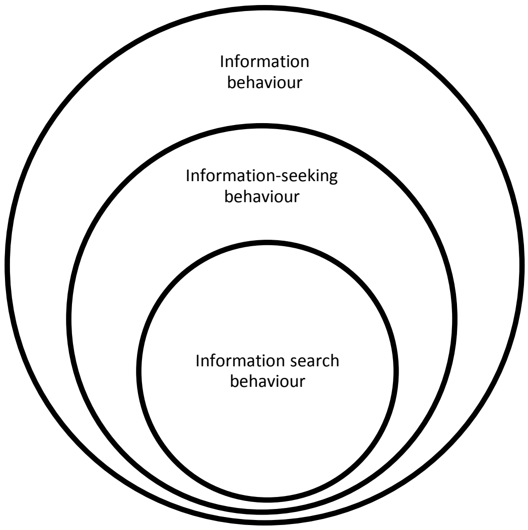


Figure 3: Wilson's (1999) Nested Model of Conceptual Areas. Reproduced with permission by Wilson.

Wilson’s model includes three conceptual areas: information behaviour, information-seeking behaviour and information search behaviour. Information behaviour is the most general concept, referring to both the activity of information seeking and the passive and purposeful behaviours related to this activity (Wilson, 1999). Activities such as communication with others and passive reception of information are included at this level.

The next level in Wilson's model is information-seeking. The main focus is purposeful seeking for information to satisfy a certain goal. At this level, individuals may interact with computer-based systems such as the Web, or manual systems such as the libraries, to resolve an information need (Wilson, 2000). Information search is subordinate to information-seeking, and refers to the actual interaction between computer-based systems and users at a micro-level. According to Wilson (2000), search is studied from different perspective on this level, for instance the interaction with the system through mouse clicking, intellectual activities, for example Boolean search strategies, and mental activities such as relevance assessment. This study is based on the innermost conceptual area in Wilson's model, namely the information search behaviour.

1.4 Methodology

There was a paradigm shift in the information seeking behaviour research in the 1980s, from a system-centred to a user-centred approach (Case, 2007). Wilson (2000) reported that a considerable number of qualitative studies on information seeking have been conducted over the last years using methods such as interviews, questionnaires and observation.

The research question addresses how dyslexia affects the interaction with search user interfaces. The attention was therefore directed at a micro-level in search logs and eye data which required a more quantitative methodology. This approach rendered possible analyses of query characteristics such as query lengths, portion of errors, time usage and eye movements (e.g. fixations and saccades). These types of data also enabled statistical comparisons between dyslexics and controls.

A triangulation of methods was applied in this study, such as screening tests, interviews and eye tracking. All participants were screened for dyslexia and tested for visual acuity. Participant data on previous search experiences, formal training and attitudes towards search systems was collected through interviews. In addition, three cognitive tests were conducted to assess short-term memory capacity and concentration skills.

The hypotheses were investigated through three experiments. The independent variable was the presence or absence of dyslexia. Participants were not assigned randomly, but were divided into one dyslexia group and one comparison group. The experiments may therefore be classified as natural experiments. Shadish, Cook and Campbell (2002) refer to natural experiments as a type of experiment where there is a naturally occurring contrast between comparison conditions. The presence or absence of dyslexia is an example of such a natural contrast which may not be manipulated.

Quantitative experiments may provide unbiased directly measurable variables which can be processed statistically. However, Case (2007) has reported that laboratory experiments may not apply in the real world. The experiments were designed to imitate well known information searching contexts. For instance, students were asked to search in databases which they commonly use. Moreover, the user search interfaces were not changed in any manner, and the tasks were considered representative for general information searching.

Search behaviour was documented through screen capturing, mouse- and keyboard input and eye tracking. Eye tracking is used to record eye movements and reported to be suitable for, among others, research on reading, scanning and visual search (Webb & Renshaw, 2008; Holmqvist et al., 2011). The eye has mainly two positions; fixations, when the eyes are still, and saccades, when the eyes move from one fixation point to another. The human eye has only full visual acuity in a small area, called the fovea. To see objects sharply, the eye has to fixate directly on the object. Consequently, eye tracking data can provide information about which areas of a computer screen that is being looked at. According to Holmqvist et al. (2011) there is also a close relation between eye-movements and cognitive processing.

According to Poole and Ball (2006), eye tracking is an unbiased method for data collecting. An eye tracker records fixations and saccades, and provides quantitative data on for instance how long a person has looked at a certain point on the screen and where the eye has moved. Based on these data, different events can be detected. For example, scanpaths can be estimated, which express the total distance the eye has travelled during one search. Scanpaths can for instance be applied in comparisons of search effectiveness between individuals. Total dwell time in Areas of Interest (AOIs) is another commonly used measurement. The experimenter may define a certain area on the display and subsequently calculate all dwells, transitions and AOI hits in this area (Holmqvist et al., 2011). Data concerning use of an AOI can for instance be used to compare the preferences for different content types on a screen.

1.5 Experiments

The three hypotheses were investigated through three experiments (see Table 1) involving 42 participants. The participants were divided into two groups: dyslexics and non-dyslexics. The experiments were conducted on a computer in a quiet and undisturbed room.

	H_1	H_2	H_3
EXP-I	✓		
EXP-II		✓	✓
EXP-III		✓	✓

Table 1: The relationship between experiments and hypotheses

The first experiment (EXP-I) concerned visual search and investigated whether graphic content can enhance search performance among dyslexics compared to searching text-only displays (H_1). The two participant groups were presented with search tasks in four conditions; icons only, words only, icons and words in a grid-layout and both modalities in a list-layout. The aim was to investigate if search performance depends on content type and identify user preferences.

The second experiment (EXP-II) explored the use of query-building aids and how misspellings affected information search in a highly tolerant search system. Both participant groups solved ten predefined search tasks each in the general Web search engine Google. This experiment was designed to address H_2 and H_3.

Experiment 3 (EXP-III) resembled EXP-II, where an equivalent set of predefined tasks were solved in the bibliographic library catalogue Bibsys Ask. This search system provides no query-building aids. Bibsys Ask has a low tolerance for spelling errors and rarely returns results for misspelled queries. The results from this experiment were compared to the results from EXP-II to provide an answer to H_3. Participants were allowed to use other information sources during the search, and in these searches several students used query-building aids in other systems. Consequently, these searches were also related to H_2.

1.6 Structure of the thesis

The thesis is structured as follows: Chapter 2 concerns theory. Topics included are universal design, dyslexia, Web accessibility and accessibility guidelines. These topics constitute the theoretical background for the research question and hypotheses. This is followed by an overview of research on information retrieval and visual search. Chapter 3 addresses methodology, such as the choice of research methods, ethics, sampling, participant variables, screening tests and experimental design. Chapter 4 summarises the research findings, and discusses the results in relation to the three hypotheses. Detailed discussions of the experimental findings are also documented in the papers attached in the appendices. Finally, the study is summarised in Chapter 5.

Six papers published in international peer-reviewed journals and conference proceedings are included in the appendices (see Appendices A – F). Appendices A and B address EXP-I on visual search. Results from EXP-II are discussed in Appendix D, while findings from EXP-III are presented in Appendix C. Consequently, the first four papers directly address the main hypotheses.

Appendix E is a summary from an icon screening and contains methodological details related to EXP-I. The paper discusses the interpretation of icons and how icons may be interpreted differently by users according to gender and age. Finally, Appendix F is a review of WCAG (W3C, 2008) to assess whether the guidelines adequately contribute to accessible Web pages for dyslexic users. This paper focuses on search user interfaces, and applies results from all the three experiments as a basis for an evaluation of each guideline in WCAG. The guidelines are also discussed in light of previous dyslexia research. This paper shows that the results from the three experiments have implications for the development of accessibility guidelines that better accommodate both dyslexic and non-dyslexic users.

Chapter 2: Background

This chapter regards the theoretical basis for the study. Section 2.1 introduces the concept of universal design, and addresses topics such as user diversity, whether universal design is an achievable goal and discusses universal design in the context of libraries. Section 2.2 introduces dyslexia, followed by Section 2.3 on Web accessibility and accessibility guidelines, both in general and targeted at dyslexic users. Section 2.4 regards information retrieval. Finally, Section 2.5 introduces visual search with a particular focus on eye movements and search performance in dual-modality displays.

2.1 Universal design

Human-computer interaction (HCI) emerged as a research field in the 1980s (Kaptelinin & Nardi, 2003), and derives from disciplines such as behavioural sciences, psychology, ergonomics and library and information science. The purpose of HCI is to understand how people use computers, and how computers can be designed to support this interaction. According to Newell and Gregor (2000), there has been an increased focus on user-centred design, user diversity and universal design in HCI during the last decades. These issues are discussed further in this section.

2.1.1 User-centred design

According to Jacob, Leggett, Myers and Pausch (1993), HCI research addresses two channels of communication. The first channel regards devices and techniques applied by computers to communicate with users, while the other includes perceptual abilities, processes and organs humans use to interact with computers. The latter refers to physical abilities such as movement, vision, hearing and touch and also cognitive abilities, for instance memory, reasoning, problem solving, skill acquisition and emotions.

In the context of HCI, the term *user* is not necessarily referring to one specific individual, but rather a group of users. Khalid (2006) emphasises that users do not constitute a homogenous group. User diversity may refer to both what differentiates people and characteristics which are common. Although most humans share some basic physiological and psychological characteristics, individual differences occur.

There are several dimensions of user diversity, such as gender, age, physical and mental abilities, education and operator skills. Differences may be long term, for instance age, gender and physical and intellectual capabilities, or short-term such as fatigue or stress. Such characteristics may also change over time. Dyslexia is an example of a long-term characteristic. However, non-dyslexics may share characteristics with dyslexics in certain situations. For instance, a person may experience reduced reading speed due to fatigue or have a higher frequency of spelling errors in stressful situations. According to Khalid (2006), understanding user diversity is valuable in the design of products.

HCI research has changed over the years, both regarding what has been studied and the methodology. This development has been described in different ways in the research literature, with terms such as phases or threads (Grudin, 2005), waves (Bødker, 2006) and paradigm shifts (Harrison, Tatar & Sengers, 2007). Although terminology and focus may differ in these models, there seems to be an agreement that the main focus of HCI has changed from systematic testing to paying more attention to contexts, emotions, cultures and experiences (Bødker, 2006).

An important concept in the development of HCI is user-centred design, which was introduced by Norman in the 1980's (Abras et al., 2004). User-centred design refers to a design philosophy and methodology based on the needs of the users. The aim is to develop and test products around user behaviour, and not demand that users change their behaviour to adapt to the products.

In the user-centred design paradigm, user requirements are considered during the whole product cycle. This is achieved through several methods, such as prototyping or usability testing. The ISO standard ISO 9241-210:2010 (ISO, 2010) is directed at requirements for user-centred design, therein referred to as human-centred design. This standard emphasises among others the importance of including users in different contexts, and throughout the whole design process.

A design paradigm closely related to user-centred design is participatory design or cooperative design, also referred to as the Scandinavian approach, where users are more actively involved as co-designers. According to Kraft and Bansler (1994) this tradition originates in Norway in the 1970s from a project where the Norwegian Iron and Metal Workers Union involved the workers in the technological development of the machines and tools they used daily at work. In contrast, the user centred design has *less direct* user involvement throughout the design process. Holone and Herstad (2013) have reported that a challenge in participatory design is to agree on how to involve people as participants. Nevertheless, there is a common belief that users should be closely integrated in the process.

While users in general have become more central in HCI research, Fuchs and Obrist (2010) have also reported an increased focus on social, ethical, political and societal implications of computer systems in the HCI literature. A topic which has been discussed is universal design, also referred to as inclusive design or design for all. Universal design originally stems from the physical realm and the architect Bednar who used this concept to describe an environment without barriers (Lid, 2013). This idea was later implemented by designers working with architecture, landscape, interior and product development, and is now incorporated into HCI research and design.

2.1.2 Universal design concept

The basic idea behind universal design is that buildings, services and systems should be available for all users without the need for any adaptation or specialised design (McGuire, Scott & Shaw, 2006). Universal design replaced the earlier practise of designing solutions directed particularly at users with impairments. For instance, in the physical realm, it was not uncommon to build two different entrances to a building, one for users who could walk stairs and an additional ramp for people who could not (Story, 1998). This is an example of accessible design, design with a purpose of meeting the requirements for people with disabilities. Such examples can also be found in computer systems. Prior to the universal design way of thought, segregated versions of Web pages for impaired users were offered for several Web sites.

According to Story (1998), accessible design is often achieved through separate design features which are often not aesthetically appealing, entails increased costs and represents a division between disabled and non-disabled users. Universal design contrasts this way of thought. Story (1998) reports that universal design is always accessible. However, it is often not noticeable, because accessibility is integrated from the beginning of the design process.

A common definition of universal design stems from the Center for Universal Design (2008):

> *«Universal design is the design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design.»*

Although several definitions of universal design exist, there seems to be an agreement about the fundamental concept. When applying universal design, user diversity is included from the beginning (Story, 1998). Universal design may be achieved by simple measures if this way of thinking has been included from the beginning in the design process, such as the barrier free entrance in Figure 4.

Figure 4 : Entrances with and without physical barriers. Photo by Tone Moseid, reproduced with permission

Iwarsson and Ståhl (2003) describe accessible design as an individual-focused approach, which specifically concerns users with impairments. A key characteristic of universal design is the focus on *all* users. It is acknowledged that people are different, but at the same time may share some common needs. Universal design is a broader concept than accessible design, and directs attention towards differences in gender, age, culture and learning preferences.

Story et al. (1998) have reported that measures which accommodate disabled users also benefit non-disabled users. Earlier, a barrier-free environment was regarded as an aim to accommodate wheelchair users. However, there might be many contexts where stairs are inaccessible, for instance for people with children in prams, delivery personnel, people with temporary impairments such as a broken leg or pregnant women with symphysiolysis. A stair-free environment would increase the accessibility for all these users.

The same principle applies to other areas, such as the production of text. For example, Santana et al. (2012) claim that by enhancing the readability of a text for dyslexic users, the text will also be easier to comprehend for people without dyslexia. Story (1998) argues that accessible design may stigmatise users with impairments and cause a feeling of segregation from the non-disabled population. Offering all users one common solution would thus reduce the discrimination of people with disabilities.

In 1998 The Center for Universal Design introduced seven principles for universal design, namely equitable use, flexibility in use, simple and intuitive use, perceptible information, tolerance for error, low physical effort, size and space for approach and use (Story, 1998). These guidelines are an example of general principles, which may be applied in several contexts, although they may be most suitable for physical environments. Separate guidelines for universal design of computer systems have been introduced, for instance the Web Content Accessibility Guidelines which addresses Web site design (see Section 2.3.1).

Formally, the concept of universal design stems from the physical realm (Lid, 2013). However, users were already central in human-centred design and participatory design, so the idea of incorporating users and user diversity was not new in HCI. According to Fuchs and Obrist (2010) the universal design perspective is coherent with a user-centred approach, and reflects a social awareness in the HCI community.

Fuchs and Obrist (2010) discuss a typology of universal design principles, but suggest a terminology related to a participatory, co-operative and sustainable design. They emphasise among others that technology should be human-centred, and thus help people in solving problems and fit their capabilities, practises and needs. The latter is closely related to universal design, and includes concepts such as perceived ease of use, where the user should be able to use a system without effort, participation in decision making, user engagement and to accommodate mental user capacities.

Newell (1995) addressed the connection between the design of computer systems for people with impairments and what he refers to as mainstream HCI studies, and proposed a division between ordinary and extra-ordinary HCI. According to Newell (1995), traditional HCI studies had regarded the user as «*an ordinary person with average abilities*». He criticised the use of a single model for the user, and introduced the concept extra-ordinary HCI. Newell (1995) applied terms such as *ordinary people* in *extra-ordinary environments*, for instance with poor lighting or much noise and *extra-ordinary users* operating in *ordinary environments*. Newell (1995) included disabled users in the term *extra-ordinary users*. Although this model implies that difficulties may arise for both people with and without impairments, the choice of words may be regarded as stigmatising. However, although the terminology may contradict the disability discourse, the thought that all users may need adaptations in certain contexts is coherent with the universal design concept.

There are different perspectives within the user-centred design paradigm, for instance mainstream design (which may often exclude disabled users), design of systems exclusively for users with disabilities and the universal design approach. Newell and Gregor (2000) have expressed scepticism towards changing the user-centred design philosophy and methodology according to universal design principles, and instead suggest the introduction of a *User Sensitive Inclusive Design*. Newell and Gregor (2000) claim that in certain situations tension may arise when including disabled users since disabled users are not in control of the research questions or evaluating the results of the research.

Newell and Gregor (2000) claim that the role of the user is to be involved in the design process without dominating it. Further, in several cases, users with disabilities may not have concurrent needs with non-disabled users. According to Newell and Gregor (2000), other issues may be problematic, such as difficulties with getting informed consent. Further, some users may not be able to express their needs while others may not be the end-users of the product. Payments may also conflict with benefit rules. Finally, some disabled users may have very specialised and little known requirements and there may be conflicting requirements between groups.

Newell and Gregor (2000) also point out methodological issues when including users with disabilities: a greater variety of user characteristics, difficulties finding representative users, possible conflicts of interest between groups, conflicts between disabled and non-disabled users and situations where it does not make sense to design for all users (for instance cars which may be driven by blind people).

Universal design is incorporated in several conventions and legislations. For instance, in the UN Convention on the Rights of Persons with Disabilities, universally designed information and communications technologies and systems is a goal (UN, 2006). Universal design is also included in EU's eAccessibility programme and in national legislations. For instance, Norwegian law demands universally designed computer systems in the Anti-Discrimination and Accessibility Act (Lovdata, 2013). This law includes an obligation to universal design of among others computer systems, including online content, and regards inaccessibility as discrimination. The Norwegian legislation (Lovdata, 2013) demands

> «*designing or accommodating the main solution with respect to the physical conditions, including information and communications technology (ICT), such that the general function of the undertaking can be used by as many people as possible*».

Regarding the Norwegian legislation, the Ministry of Children, Equality and Social Inclusion (2010) realised that a lack of formal regulations may cause difficulties in enforcing the law. Consequently, a regulation on ICT was added (Kommunal- og moderniseringsdepartementet, 2013). This regulation demands among others coherence with Web Content Accessibility Guidelines 2.0 (W3C, 2008) or an equivalent standard. WCAG, discussed further in Section 2.3.1, is also incorporated in the national legislation of several countries, such as Australia (Australian Human Rights Commission, 2014), Canada (Treasury Board of Canada, 2013) and the EU (The European Parliament, 2002).

2.1.3 Design for all?

Newell and Gregor (2000) suggested the term *User Sensitive Inclusive Design* instead of altering the user centred design paradigm to incorporate universal design. One of the reasons behind this suggestion was to express what Newell and Gregor (2000) regard as a more achievable goal than to design for universality. This issue has been discussed by researchers and developers. Is it possible to design a product which may be usable for all people?

Newell and Gregor (2000) emphasise conflicting needs with unimpaired users or between users with different disabilities as a counter-argument against universal design. They also argue that in some cases it makes no sense to make a universally designed product due to the nature of the product. Examples of this may be motorcycles or cars that may be operated by blind users or sound operating systems for deaf users. Newell and Gregor (2000) argue that aiming at an unachievable goal «*has the danger of inhibiting people from attacking the problem at all*».

Harper (2007) has suggested to apply the term *Design-for-one*. Harper (2007) argued that within the universal design perspective there is a need to make generalisations about users, which may lead to the exclusion of certain individuals who do not fit into these generalisations. Further, by addressing all users in one design, there is a risk that the product will not be usable for anyone. Harper (2007) claims that universal design should be discussed by focusing on the systems and the method of controlling the system rather than on the user interface itself. Further, Harper (2007) argues that interfaces should be designed for a target user group with opportunities for adapting the system according to personal interaction requirements. Consequently, *design for one* is suggested as a counterweight to *design for all*, which Harper (2007) regards as impossible to achieve.

In some cases, it may be difficult to accomplish one accessible solution for all due to, among others, differing or conflicting needs. For instance, blind users may rely on auditory content while deaf users might prefer visual content. Further, some users may need individual adaptation, assistive technology or a personal assistant to be able to utilise a product. According to the definition applied by the Center for Universal Design (2008), products and services should be accessible «*by all people, to the greatest extent possible*». In contrast, the Norwegian Anti-Discrimination and Accessibility Act (Lovdata, 2013) applies the phrase «*as many people as possible*». This may be regarded as a less strict (and possibly more realistic) definition, which accepts that it might not always be possible to accommodate a whole population through one design.

To illustrate the issues with universality and accessibility, the National Resource Centre for Accessibility and Social Inclusion in Norway proposed the usability pyramid (see Figure 5). Although the term usability is applied in this model, the main focus is accessibility. While usability regards effective and efficient design, accessibility concerns creating equivalent user experiences for disabled users.

According to the usability pyramid, universal design is the main strategy for inclusion (Universell, 2008), and should therefore be the overall goal. However, the purpose of the model is to show that universal design may not be enough to ensure accessibility for all people. In certain contexts, there is a need for adaptation or assistive technology, measures which constitute the remaining part of the usability pyramid. There is a proportional relationship between the four levels in the model. By increasing the size of one level, the remaining levels decrease. Consequently, investing more money and effort into universal design may reduce the need for further adaptation for both groups of users and individuals.

The usability pyramid may be applied for instance to analyse a certain service to identify different strategies related to user diversity and disabilities. For instance, Berget and Moseid (2011) applied the model to investigate the accessibility of Norwegian public library facilities. Libraries that are physically universally designed will increase the accessibility for all, and reduce the need for measures at the remaining levels. Nevertheless, the library may still not be accessible for some users, and additional initiatives must be implemented to ensure access for all.

An example of level two measures, adaptation for user groups with concurrent needs, is the Norwegian Library of Talking Books and Braille (NLB), which offers library services to users who have visual impairments or reading disorders. Another example is the inclusion of books containing sign language or Braille in the library collection. The third level, individual adaptation and assistive technology, includes services such as volunteers who reads aloud for users who cannot read themselves, for instance elderly or cognitively impaired users. The fourth level regards personal assistance, which is adapted to each individual user, for instance for people with severe autism.

The example above illustrates how the usability pyramid may be used to render visible that universal design may not always be adequate to ensure accessible services for all people in a society. For certain users and in particular contexts, measures on different levels may be necessary. However, the model also illustrates that the need for level two, three and four measures may be reduced if more resources is put into universal design measures which will include more users.

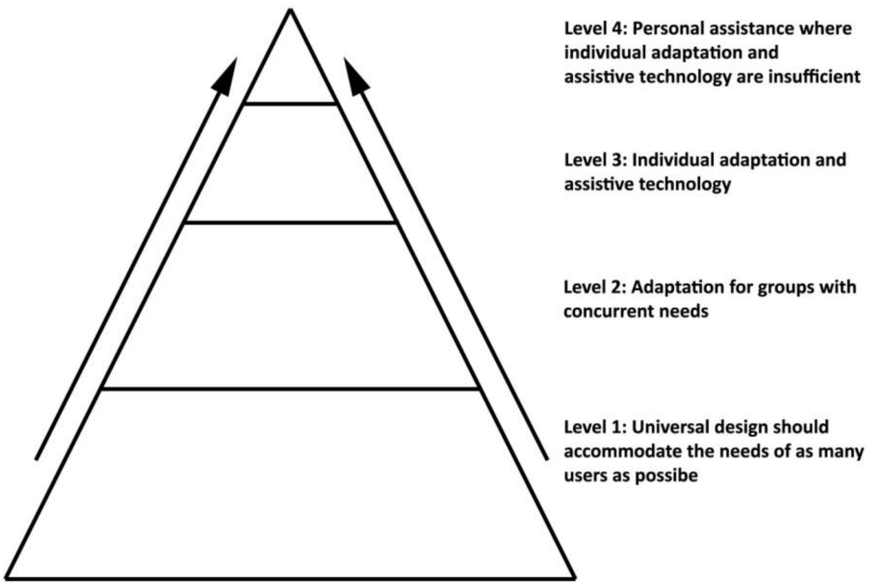


Figure 5: The usability pyramid (translated reproduction based on a figure in Universell, 2008)

2.1.4 Universally designed libraries

In the context of universally designed libraries, the physical environment, the library collection and the electronic services must be considered. According to Schmetzke (2006), concerns for accessible information and electronic resources originally caught on within the field of library and information science, although the focus was initially on adaptive technology. An example of the initial perspective is the paper by Cantor (1996), who focused on the inclusion of assistive technology in the libraries, and discussed which equipment libraries should provide, for instance speech synthesis, magnifiers and wheelchairs. Accessible Web design got widely coverage in the library literature from 1996, when among others Paciello (one of the persons who were instrumental in the Web Accessibility Initiative in W3C) published several articles on accessible Web design in library journals Schmetzke (2006).

In 2004, the Museums, Library and Archives Council in UK commissioned a Web accessibility audit that included user testing and automatic testing and was an attempt to investigate the accessibility of libraries, museums and archives (Weisen, Petrie, King & Hamilton, 2005). Weisen et al. (2005) reported that academic libraries had fewer violations than public and specialist libraries. However, the report concluded that only a small number of Web sites complied with accessibility guidelines.

The user testing involved blind users, dyslexics and partially sighted. Dyslexic users were able to complete 82.5% of the tasks, compared to 77.5% for partially sighted and 66.7% for blind users. The five most frequent problems identified were related to page design, ambiguously named links, colour scheme and contrast, no “skip navigation” link and difficulties with external links. The report did not discuss search user interfaces. Consequently, it seems like only the library Web pages were included in the audit, and that library catalogues were not investigated.

Burgstahler (2002) discusses online distance education in the context of universal design, and emphasises the need for accessible library resources. Without universally designed library research resources, including catalogues and indexes to journals, Burgstahler (2002) argues that distant learning becomes inaccessible for a number of students, among other users with reading disabilities.

In 2001 the Norwegian Archive, Library and Museum Authority initiated the project «*The accessible library*», with the aim to develop more accessible libraries for disabled users. Three public libraries were used for a pilot project with a focus on three main issues: physical accessibility, the presence of an adviser with impairments and the effect of user participation. A list of 114 measures to obtain a more accessible library was developed (Norwegian Archive, Library and Museum Authority, 2004), and addressed topics such as parking spaces, stair-free entrances and the distance between bookshelves.

Only two guidelines addressed technology, namely that there should be computers with Internet access that are accessible for users with physical or visual impairments. Consequently, there was little focus on the digital library, and search systems were not addressed at all. Information technology was mentioned on the project Web site, including accessible public computers, Web sites and audiobooks (The accessible library, n.d.). However, the work related to more accessible Web pages did not address search systems, and were not included in the aforementioned guidelines.

An important part of the library is the collection, and audiobooks are now commonly found in most libraries. These are examples of books which are accessible for many different users, such as people with reduced vision, reading disabilities or other learning disabilities and illiterate library users (Moyer, 2012), but are also popular by non-disabled users. Offering an extensive collection of audiobooks may therefore be an important step towards a universally designed library collection, especially if the books are marked with braille, so that they can easily be found by visually impaired users as well. Libraries often also offer books with large letters, braille or sign language. Such books are examples of adaptation to certain user groups, which may increase the overall library services for a diversity of people.

Special libraries, such as the Norwegian Library of Talking Books and Braille (NLB), offers audiobooks and braille books for users with visual impairments and reading disabilities. These libraries do not represent universal design, since they do not allow users outside the defined user group access to their collections, but merely focus on a small group of users. Universal design involves several user groups, and does not only address disabled users. Ciszek (2012) argues that academic libraries must pay attention to several underrepresented groups, and draw attention to multiculturalism, LGBT (Lesbian, Gay, Bisexual and Transsexual) students and disabled students.

2.2 Dyslexia

Different user groups may share some common characteristics, and universal design should benefit a majority of users regardless of gender, culture, ages and functional levels. McCarty and Swierenga (2010) reviewed existing research on dyslexia, and reported that a common finding was that actions which make services more accessible for dyslexic users also result in more accessible products for non-dyslexics. Dyslexia is the only controlled variable in this study, which includes an equally sized user group of dyslexics and non-dyslexics. The overall purpose is to improve search user interfaces for both groups. Consequently, the study should comply well within the universal design domain.

2.2.1 Defining dyslexia

Dyslexia is a learning disorder which mainly affects reading and writing development (Snowling, 2000). The word dyslexia stems from the Greek words *dys* (impaired) and *lexis* (speech), and was first introduced by Rudolf Berlin in 1887 to describe the loss of reading abilities (Smythe, 2011). Gallagher, Frith and Snowling (2000) have reported that dyslexia is highly hereditary, and according to Hawke, Olson, Willcut, Wadsworth and DeFries (2009) it is overrepresented in the male population.

The prevalence of dyslexia is estimated to 5-10% of any population (Ahissar, 2007). However, dyslexia is found in different degrees, and varies according to language. For instance, Ziegler and Goswami (2005) reported a higher prevalence in languages with a low consistency between orthographic and phonologic units, also referred to as low spelling to sound consistency. Norwegian orthography is regarded as a semi-transparent language (Lyster, 2007) with 29 letters and 36 graphemes representing 40 phonemes (Helland & Kaasa, 2005). In contrast, English orthography consists of 26 letters and 41-44 phonemes which can be written in 561 manners. According to Helland and Kaasa (2005) there is a lower spelling to sound consistency in English compared to Norwegian, which makes English language more difficult for dyslexics.

Norwegian children are reported to learn to read and write quite easily (Lyster, 2007). Nevertheless, dyslexia is still considered a major challenge in Norwegian education. Typical errors among dyslexics relate to compound words: *fredspris*, double consonants: *Oppland*, silent consonants: *Sigrid* (d is not pronounced), consonant clusters: *Undset*, silent vowels: *Algerie* (e is not pronounced) and words with irregular orthography, where the spelling and pronunciation differs: *Shakespeare* (Helland & Kaasa, 2005; Lyster, 2007; Høien & Lundberg, 2012), all of which are common in the Norwegian language.

Francks, MacPhie and Monaco (2002) claim that the dyslexia profile varies in physiological, behavioural and cognitive measures. Dyslexia is chronic and persistent, and remains throughout life (Shaywitz & Shaywitz, 2005). Many people may not be aware of their dyslexia (Mortimore & Crozier, 2006). However, according to Francks et al. (2002), dyslexia may be traced in small children before they have started reading, for instance through delayed acquisition of speech and a high frequency of speech errors. Diagnoses are based on various tests such as word recognition, orthographic coding, rapid automated naming, phonological awareness, phonological decoding and spelling and working memory tests (Francks et al. 2002; Helland, Tjus, Hovden, Ofte & Heimann, 2011).

Although the main characteristics of dyslexia are agreed upon, the definition of dyslexia is much debated. According to the classic definition, dyslexia is a specific learning disorder where achievement levels of reading are significantly below the expected level, given general intelligence (Snowling, Duff, Petrou, Schieffledrin, & Bailey, 2011). However, the focus has shifted from intelligence levels to phonological processing.

Høien and Lundberg (2012) have suggested defining dyslexia as «*a disturbance in dealing with the code of the written language based on a deficit in the phonological system of the spoken language*». This is in accordance with the definition applied by the International Dyslexia Association (2016):

> *«Dyslexia is a specific learning disability that is neurobiological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede growth of vocabulary and background knowledge.»*

According to Snowling (2000) there are different variations of dyslexia. For instance, some dyslexics are slow but accurate readers, while others are fast but inaccurate readers (van der Schoot, Licht, Horsley & Sergeant, 2000). Nevertheless, difficulties with word decoding and reading comprehension are commonly found in both types of readers (Shaywitz & Shaywitz, 2005). A basic component of reading is to identify phonemes and combine them into words. This is often referred to as phonological awareness, and is reported to be missing in children, adolescents and adults with dyslexia (Shaywitz & Shaywitz, 2001). Although students may become more accurate readers over time, Shaywitz and Shaywitz (2001) claim that reading will remain slower and less automatic compared to their peers.

The aetiology of dyslexia is disputed (Jones, Branigan & Kelly, 2007; Schuchardt, Maehler & Hasselhorn, 2008). Most researchers agree on an underlying phonological deficit (Snowling, 2001), which affects the processing and representation of speech sounds. However, some researchers regard these phonological deficits as secondary, and discuss a more auditory basic impairment, or suggest that dyslexia may be a part of a sensorimotor deficit, also referred to as the magnocellular theory (Ramus, 2003). It is also debated whether several subtypes of dyslexia exist, such as surface, phonological, mixed and mild dyslexia (Ziegler, Castel, Pech-Georgel, George, Alario & Perry, 2008).

Comorbidity between dyslexia and other diagnoses is not uncommon. For instance, according to Germanò, Gagliano and Curatolo (2010), 18-20% of dyslexics may also have Attention Deficit Hyperactivity Disorder (ADHD) or Attention Deficit Disorder (ADD). ADHD and ADD are typically characterised by excessive activity, impulsivity and a short attention span, which refers to the time a person can concentrate on a task without being distracted (Germanò et al., 2010). Prevalence between dyslexia and other diagnoses are also suggested, such as the mathematical impairment dyscalculia (Landerl, Fussenegger, Moll & Willburger, 2009), the writing disorder dysgraphia (Nicolson & Fawcett, 2011) and dyspraxia, an impairment in the organization of movement (Gibbs, Appleton & Appleton, 2007).

2.2.2 Cognitive profile

Although researchers do not agree upon how to properly define dyslexia, there is more consensus regarding the cognitive profile. Several characteristics have been typically associated with dyslexia, such as reduced short-term memory capacity, impaired rapid naming skills, concentration difficulties and low self-esteem (Snowling, 2000; Snowling, 2001; Caroll & Iles, 2006; Mortimore & Crozier, 2006; Smith-Spark & Fisk, 2007). These characteristics may have implications for experimental design.

A reduced short-term memory capacity is regarded as one of the most prominent characteristics of dyslexia (Smith-Spark & Fisk, 2007). Short-term memory is the part of the brain system which provides temporary storage and manipulation of the information that is needed for a variety of complex cognitive tasks, such as language comprehension, reasoning and learning (Baddeley, 1992). Short-term memory is also suggested to play an important part in prose comprehension and verbal reasoning (Baddeley & Hitch, 2008). Short-term memory has limited capacity. Miller (1956) suggested a limit of seven, plus or minus two. However, this limit has been modified in later research, for instance by Cowan (2008), who suggested that the capacity limit is typically three or four units, and that more items cannot be stored as separate items.

Short-term memory may be separated into visual-spatial memory and verbal memory. Visual-spatial memory concerns storage of information about the environment. It is suggested that this part of the short-term memory has two components, one which stores visual information such as colour and shape and one part that stores, among others, spatial placement (Mammarella, Pazzaglia & Cornoldi, 2008). Visual-spatial memory is typically limited to three or four objects (Baddeley, 2003), and is measured by tests such as The Corsi Block Tapping-test (see Section 3.3.3). Verbal memory stores information related to verbal-linguistic information, and is believed to predict measures of text comprehension (Kane et al., 2004). According to Baddeley (2003), this part of the memory is often related to the phonological loop, plays a part in the acquisition of foreign languages and is assumed to be impaired in children with specific language disabilities. Verbal short-term memory can, among others, be tested through a Digit Span test (discussed further in Section 3.3.4).

According to Siegel and Ryan (1989) short-term memory is regarded as an important component of reading, for decoding words, reading sentences and comprehending longer texts. For instance, information about word meaning, syntax and phonological rules must be retrieved while words and sentences must be processed before meaning is comprehended. Daneman and Carpenter (1980) emphasise that another important feature of reading is to remember the preceding text and recall it as a context for subsequent text to create meaning.

Smith-Spark and Fisk (2007) have reported that impaired short-term memory is well documented in dyslexics, and that this impairment extends into adulthood. The reduced short-term memory capacity has certain implications for experimental design. For instance, instructions should be short and concise. Further, tasks should be designed so that they appear essentially different from previous tasks, thus avoiding that participants mix up tasks.

Reduced rapid naming skills are commonly associated with dyslexia. Rapid naming skills refers to the ability to connect visual and verbal information by assigning appropriate names to for instance objects, colours or numbers (Lervåg & Hulme, 2009). According to Neuhaus, Foorman, Francis and Carlson (2001), rapid naming skills are assumed to affect reading because it relates to, among others, word-recognition and letter-naming. Catts et al. (2002) have reported a general deficit in processing speed among people with reading difficulties. Lervåg and Hulme (2009) suggest that achievement levels in naming pictures and letters can be used as predictors for reading abilities among pre-literate children.

In this study it was assumed that reduced rapid naming skills may affect the abilities to transfer information needs into actual queries, and may thus affect time usage per query and the choice of query terms. Consequently, reduced rapid naming skills were considered when formulating the search tasks for the information retrieval experiments (EXP-II and EXP-III). This issue is discussed in more detail in Section 3.4.

According to Jeffries and Everatt (2004), several dyslexics struggle with mental organisation. Concentration difficulties are also reported to be common, especially when focusing on a task for a long time (Mortimore & Crozier, 2006). Consequently, there is a trade-off in experimental design between experiment length and comprehensiveness. If experiments are too long, a fatigue-related effect may introduce a bias.

It has been reported that dyslexia affects self-concept and self-esteem. Children with dyslexia may feel isolated and excluded at school (Humphrey & Mullins, 2002), and low self-esteem issues and a fear of failure among dyslexic students have been well documented according to Caroll and Iles (2006).

Dyslexic users may feel discomfort related to testing situations and display performance anxiety during task solving. Consequently, experiments should be designed and presented in a manner which minimises this anxiety. For instance, tasks may be presented with increasing difficulty levels, so that dyslexics do not despond and terminate the experiment at an early stage. It may also prove helpful to include some simple tasks subsequent to difficult tasks. Nielsen and Levy (1994) claim that users tend to blame themselves rather than the system when failing to solve a task. Consequently, it should be emphasised during instructions that the main object for testing is the system, not the users.

Sutton, Erlen, Glad and Siminoff (2003) have emphasised that recruiting volunteers from vulnerable populations is difficult. This finding may also apply to dyslexic students, since low self-esteem issues and performance anxiety may reduce the willingness to participate.

2.3 Web accessibility

A basic principle of universal design is the consideration of different human factors, such as vision, hearing, speech, body function and cognition. In the context of cognition, several human abilities are suggested to be important when designing user interfaces. Story et al. (1998) have emphasised concentration, comprehending visual information, understanding or expressing language, performing operations in the proper order (sequencing), organisation and memory capacity (both short and long term). According to the cognitive profile of dyslexia, all these characteristics may be impaired in dyslexic users.

Story et al. (1998) introduced a checklist to accommodate users with varying cognition. For instance, can the system be used if you cannot read?, must operations be performed in a certain order?, can the system be used slower than intended? and what happens if the user makes a mistake or is distracted while using the system?. Principles and guidelines for universal design have been introduced to address such questions. The most prominent and widely used guidelines addressing Web content the the Web Content Accessibility Guidelines, also referred to as WCAG (W3C, 2008).

2.3.1 WCAG

Several Web accessibility guidelines exist. Such guidelines may be directed at the general population and include a wide range of disabilities, such as WCAG (W3C, 2008), or focus on certain groups, for instance visually impaired users (Leuthold, Bargas-Avila & Opwis, 2008) or dyslexics (Santana et al., 2012). According to Moreno, Martínez and Ruiz-Mezcua (2008), WCAG has become the de facto standard applied by developers, designers and legislators.

The first version of WCAG 1.0 was developed by the World Wide Web Consortium (W3C), and was endorsed as a W3C Recommendation in 1999 with a primary goal to promote accessibility (W3C, 1999). These guidelines were in 2008 replaced by WCAG 2.0 (W3C, 2008), which is also defined as an international standard, ISO/IEC 40500:2012 (ISO, 2012).

WCAG is based on four principles, namely that Web content should be perceivable, operable, understandable and robust. A total of 12 guidelines, one to four per principle, are formulated. These guidelines are divided into more specific success criteria, which are intended to be testable and regarded as platform independent (W3C, 2008). Each criterion is categorised according to a conformance level, where A is the lowest and AAA the highest. Several success criteria are quite general and may apply to several user groups, others address specific impairments. Dyslexia is not specifically mentioned in WCAG, but is included in broader terms such as «*cognitive, language and learning disabilities*» and «*print disabilities*» (W3C, 2008).

In the context of dyslexia, a few success criteria in WCAG 2.0 may be particularly important. For instance, guideline 2.2 under the principle of operability concerns providing the users enough time to read and use content (W3C, 2008). This guideline is important for dyslexics who, according to van der Schoot et al. (2000), are either slow readers or fast inaccurate readers, who may need to re-read passages to comprehend meaning. Enough time is also important when performing operations such as filling out forms, ordering tickets or paying online. It has been reported that dyslexics struggle with unclear navigation (Disability Rights Commission, 2004; Al-Wabil et al., 2007). Guideline 2.4 addresses navigable content, and stresses that there should be mechanisms to help users navigate, find content and determine where they are (W3C, 2008).

Principle 3 regards understandable interfaces. Guideline 3.1 states that the reading level should be in accordance with lower secondary educational level. This is the criteria which may be clearest targeted at users with reading impairments. However, clearer text is also reported to improve the usability for non-dyslexics and users with other impairments, such as visually impaired (Evett & Brown, 2005) or deaf users (Boldyreff, Burd, Donkin & Marshall, 2001). Included in the third principle are also guidelines concerning predictable navigation (3.2), automatic detection of input errors and error suggestions (3.3). The latter may be particularly helpful for dyslexics, since this group is reported to commonly produce spelling errors (Snowling, 2000).

2.3.2 Web accessibility and dyslexia

According to McCarthy and Swierenga (2010), dyslexics are rarely discussed thoroughly in the context of Web accessibility. Dyslexia research regarding accessibility is often directed at general topics, such as clear text (Evett & Brown, 2005), font (Rello & Baeza-Yates, 2013; Rello, Pielot, Marcos, & Carlini, 2013) and Web navigation (Al-Wabil et al., 2007).

In the context of dyslexia, several researchers have addressed fonts, both regarding typeface, size and layout. According to Rello and Baeza-Yates (2013), font types have been found to affect reading performance. It has been suggested that dyslexics prefer font types without serifs (Evett & Brown, 2005), and that italics should be avoided (Rello & Baeza-Yates, 2013). Font size has also been reported to have a significant effect on readability (Evett & Brown, 2005; Rello et al., 2013). Regarding layout, Zorzi et al. (2012) found that increasing letter-spacing improved reading performance among Italian and French dyslexic children, because the increased spacing reduced the effect of crowding (see Section 2.5.2). However, Rello et al. (2013) concluded that line spacing had no significant effect on reading performance.

Special fonts for dyslexic users have also been designed, such as Sylexiad, developed for adult dyslexic readers (Hillier, 2008), the font Dyslexie and OpenDyslexic. Rello and Baeza-Yates (2013) included the type OpenDyslexic in their study on reading performance, but found no significant differences compared to other fonts. However, Rello and Baeza-Yates (2013) suggested that a lack of reading aloud tests as an explanation for this finding, since these fonts are designed to improve reading aloud performance.

Researchers have also addressed online text production by dyslexics. Baeza-Yates and Rello (2011) estimated the amount of texts written by dyslexics on the Web based on the presence of typical dyslexia-related errors (discussed in Section 2.2.1). Baeza-Yates and Rello (2011) found that dyslexics were underrepresented as content producers, and that the implementation of natural language processing tools could improve Web accessibility for dyslexic users.

Regarding colours, Kurniawan and Conroy (2007) found increased reading speed for dyslexic and non-dyslexic readers when they were allowed to select colour scheme. This is in accordance with guideline 1.4.3 of WCAG, where the instruction states that sometimes it may be advantageous to allow users to view content in their preferred colour schemes (W3C, 2015).

On a usability test concerning Web pages, The Disability Rights Commission (2004) found that dyslexic users failed on 17% of the tasks. Five main problems were identified: unclear and confusing layout of pages, confusing and disorienting navigation mechanisms, inappropriate use of colour and poor contrast between content and background, graphics and text too small, and complicated language or terminology (Disability Rights Commission, 2004).

Al-Wabil et al. (2007) interviewed ten dyslexics to investigate how dyslexia affects Web navigation. They found that dyslexics experienced significant navigational barriers online, did not find internal search (the search function on a Website) useful, and mainly used internal search as a last resort when navigation failed. Dyslexic users did not approve of the quality of the results, and experienced difficulties when spelling errors were not accounted for. One dyslexic used Microsoft Word for spell checking before including a term in a search box, while another student commented on using Google for searching information on the university Web pages due to an inaccessible internal search system (Al-Wabil et al., 2007). Studies addressing the information retrieval of dyslexics are introduced in Section 2.4.2.

WCAG 2.0 is directed at users of all ages and considers various impairments. However, W3C (2008) addresses difficulties concerning such a wide range of users in the introduction to their guidelines: «*Note that even content that conforms at the highest level (AAA) will not be accessible to individuals with all types, degrees, or combinations of disability, particularly in the cognitive language and learning areas*»

Until WCAG provides sufficient guidelines for all user groups, there may be a need for additional guidelines directly targeted at specific groups, such as dyslexic users. Friedman and Bryen (2007) reviewed 20 guidelines for people with cognitive impairments where three of these specifically addressed dyslexia. They found that the most frequently cited guideline regarded the use of pictures, icons and symbols in addition to text, which was mentioned in 75% of the guidelines. Other issues addressed were clear and simple text (60%) and consistent design (60%). The two latter guidelines seem to be in accordance with WCAG (W3C, 2008). None of the 22 most discussed issues regarded search user interfaces, except for a point to alert users on possible errors.

Santana et al. (2012) reviewed and summarised 41 dyslexia-specific guidelines. These guidelines covered navigation, colours, text presentation, writing, layout, images and charts, end user customization, mark-up, videos and audios. In the context of information searching, guideline 5.1.6, regarding internal search is of particular interest. Santana et al. (2012) emphasise the need for writing aids, a topic which was explored in the two information retrieval experiments. Further, guideline 5.6 states that images and pictures should complement textual information, and that images and icons should be particularly used in links. Santana et al. (2012) claim that dyslexics tend to consider images over words. This issue was investigated in the visual search experiment.

The British Dyslexia Association (n.d.) has also developed a style guide consisting of three parts: dyslexia friendly text, accessible formats and Website design. The first two sections include general advice on issues such as font types and sizes, and emphasise short and simple text. The part on Websites emphasises clear navigation and the use of graphics, images and pictures to break up text. Zarach (2002) has presented a similar set of ten guidelines targeted at dyslexic users. These guidelines recommend customisation of colours, using images alongside text, keeping text and navigation simple, and finally addresses the use of sans-serif fonts and a recommended font size.

Existing guidelines directed at dyslexics seem to focus on approximately the same issues, and it does not seem to add much new elements to apply several different guidelines. Further, most of the topics addressed seem to be included in WCAG, particularly colours, navigation and simple text. Fonts and images, however, are not addressed in WCAG, neither is the design of search user interfaces.

2.3.3 Assessing Web accessibility

Designing for a wide spectre of users may be difficult. However, assessing Web sites for accessibility may be a useful method to reveal shortcomings in user interfaces and functions. Several evaluation methods exist, such as expert evaluation and user testing. Heuristic evaluation is a commonly used evaluation method involving one or several experts who compares a set of criteria to for instance a Web site or computer programme with an aim to identify discrepancies. According to Nielsen (1994), this method is relatively fast, easy and inexpensive, but should be completed by more than one expert, since one person rarely discovers all usability problems in one interface.

Nielsen (1994) introduced a widely used list of general heuristics: visibility of system status, match between system and the real world, user control and freedom, consistency and standards, error prevention, recognition rather than recall, flexibility and efficiency of use, aesthetic and minimalistic design, help users recognize, diagnose and recover from errors, in addition to help and documentation. In the context of search systems, error prevention relates to the tolerance level for misspellings, and is addressed by H_3 and the two information retrieval experiments.

Some of the criteria in Nielsen's (1994) heuristics may accommodate users with dyslexia, such as considering the use of words and reducing memory load. However, more extensive accessibility guidelines, such as WCAG (W3C, 2008), may reveal supplementary shortcomings in the interface. Mankoff, Fait and Tran (2005) have reported that a few accessibility criteria can be tested by automated tools, but several guidelines must be manually tested. For instance, an automated system may confirm whether alternative text is added to an image, which may increase the accessibility for visually impaired users. However, such a system may not judge if this text is helpful to describe the image for users who cannot see it, or whether it includes less useful text, such as «*this photo is taken by person NN*» or «*Text to be added here*».

A diversity of users may perform a multitude of tasks using one interface or system. Nielsen (1994) claims that evaluators cannot be expected to have knowledge of all of these users. Further, lack of experience with disabilities or the use of assistive technology may prove problematic with the heuristic method. Consequently, Petrie and Bevan (2009) proposed that heuristic evaluation is insufficient for accessibility testing, and that actual user testing is required.

Evaluation involving users is common within user-centred design. According to Petrie and Bevan (2009), it is preferable that users participate in the whole process. However, since user testing may be an expensive method, one may first conduct a heuristic evaluation eliminating obvious design errors and system failures followed by user testing at a later stage. In some cases, certain problems may be overlooked in user testing. Nielsen (1994) therefore recommends both methods to be applied.

A potential shortcoming with heuristic evaluation is that if the heuristics do not accommodate all users or needs, the evaluation may overlook important problems. This is in accordance with Richards and Hanson (2004) and Rømen and Svanæs (2012), who found that although Web pages comply with WCAG, this is not sufficient to guarantee accessibility. This is supported by Leuthold et al. (2008) who found that compliance with WCAG did not make any difference in Website usage for people with visual impairments. W3C (2008) has pointed out potential shortcomings in WCAG for users with impairments related to cognition, language or learning. Consequently, heuristic evaluation of user interfaces, such as search systems, may not be sufficient when assessing accessibility for dyslexic users.

User studies often involve one or more persons in one or several stages during development, using methods such as interviews, prototyping and observation (Spinuzzi, 2005). The user may be in a lab or work remotely, and feedback may be given in different ways, such as thinking aloud or taking notes. User behaviour may be documented by for instance screen capturing or eye tracking (Ehmke & Wilson, 2007). It has been discussed how many people who are needed in a successful user test. Petrie and Bevan (2009) suggest including eight people, or even more if there are distinct target users.

Petrie, Hamilton, King and Pavan (2006) claim that recruiting difficulties are well known for user tests which includes disabled or old users. Burgstahler and Doe (2004) have pointed out that recruiting issues are sometimes solved by abled-bodied users who simulate various disabilities, but emphasise that this method is regarded by many people as quite controversial. The ethics of this method is also debated. Another alternative is to include especially trained users to test for several disabilities

McCarthy and Swierenga (2010) have suggested that increasing the accessibility of a Web site for dyslexic users will also increase accessibility for users without dyslexia. Consequently, if user studies include only a few people, it may be expedient to include at least one dyslexic. This is supported by Dixon (2007), who found that dyslexic students identified more general issues and provided more detailed feedback than non-dyslexics. This view is also supported by Rupprecht, Blum and Bomsdorf (2014), who concluded that dyslexics make an essential contribution to user testing, and that such users should be involved in an early stage of the design process.

2.4 Information retrieval

Although HCI practitioners have directed much attention to Web page design, Zhang and Salvendy (2001) claim that the perspective may be limited regarding cognitive abilities in the context of information retrieval performance. However, the interaction between users and search systems has received much attention within library and information science (LIS). Wilson (1999) divided this research into three areas in his Nested Model of Conceptual Areas (see Figure 3): information behaviour, information seeking behaviour and information search behaviour. Typical topics for discussion have been how information needs arise, how users search for information and how information is evaluated and applied. As discussed in Section 1.3, the experiments conducted as a part of this study regards information searching at a micro level, which is represented by the innermost level in Wilson's model.

H_2 and H_3 regard search behaviour in two existing search systems, namely a general Web search engine (EXP-II) and a library catalogue (EXP-III). This section provides a general overview of information seeking research, with an emphasis on information searching. Although a large amount of theories and models exist, Beverly, Bath and Barber (2007) have reported that none of these seem to incorporate disabilities. Nevertheless, several existing theories may have transfer value to users with disabilities in general or dyslexics more specifically.

2.4.1 Information needs and query formulation

Searching for information may be a mentally challenging task. An information need, which is often vague and unclear, must be converted into an actual query and inputted into a search box. Further, the query has to include words which are comparable with the search system's database. Overall, information search requires a significant amount of cognitive load. Gwizdka (2010) has pointed out that query formulation is the most demanding task during search. An aim for search user interfaces should therefore be to assist the users and thus reduce the cognitive load demanded for successful information retrieval.

Many theories concern how information needs arise and are resolved. Although the terminology varies, a common feature in several models is uncertainty. For instance, Belkin (1980) claims that an anomalous state of knowledge (ASK) arises when there is an anomaly between the user's state of knowledge related to a topic or situation. According to Belkin (1980), an information need arises when the user recognises the anomaly. However, the user is unable to exactly specify what is necessary to solve the ASK. As a consequence, Belkin suggests that information retrieval systems should let the user describe the ASK rather than to formulate a specific query.

Dervin (2005) developed a theory called Sense-Making, where users move in time and space until they encounter a cognitive gap where knowledge is needed to solve a situation and thus bridge the gap. The primary goal with the information seeking, Dervin (2005) claims, is thus to make sense of a situation. Kuhlthau (1991) included feelings and thoughts in her model on information seeking and divided information seeking into a six-stage process, from initiation to presentation. The first three phases include uncertainty, confusion and frustration, which are transformed into clarity and sense of direction as the searching progresses.

Expressing an information need may be difficult because, among others, the user must formulate a query concerning an unfamiliar topic. Taylor (1968) introduced a four-step process on how an information need is transferred to an actual query, also referred to as Question-Negotiation. The first phase involves a visceral need, which is vague and unexpressed, that is transformed into a conscious need. The information need is then developed into a formalised need, which is expressed by the user with actual words. The last phase is a compromised need, which is the query submitted to the search system.

According to Taylor (1968), the conversion from a formal need to a compromised need may, among others, be resolved through the assistance of librarians. An important aspect of Taylor's theories is that queries may be difficult to formulate, and in certain cases the user needs assistance from either librarians or the search system. The early stages in the query-negotiation may be compared with Belkin's ASK-model (1980), where the information need is considered vague and difficult to express.

Regardless of which theories or models on information seeking that are applied, a common conception seems to be that searching is difficult, and that users may become frustrated when converting information needs into actual queries. Hearst (2009) has pointed out that the function of a search system is to assist the users regardless of age, culture and background. Hearst (2009) has suggested eight design criteria for search systems: reduce short-term memory load, reduce errors, offer informative feedback, support user control, provide shortcuts, consistency, permit easy reversal of actions and design for closure. The first two criteria may be particularly important for dyslexic users, who are reported to have a reduced short-term memory capacity (Smith-Spark & Fisk, 2007) and frequently produce spelling errors (Snowling, 2000). The second criterion is addressed by H_3.

Most search systems provide at least two search user interface options; basic and advanced search. Basic search usually consists of one search box in which the whole query is inputted. In contrast, advanced search typically has several boxes which require different input types, such as titles, author names or keywords. Certain search user interfaces also provide field search, where queries are formulated by manually inputting query terms and field codes in one large search box. Boolean operators, such as AND, OR and NOT, are usually applicable in all the three search user interface options. According to White and Morris (2007), the use of such advanced search options correlates with the level of expertise, and novice users rarely apply advanced search techniques.

Search systems vary according to input requirements and search options. Some systems provide full-text search, where the user may formulate queries using natural language. This is common in Web search user interfaces (Hearst, 2009). In contrast, other systems match the queries against document surrogates, containing metadata such as author name, titles and keywords, and in some cases searchable abstracts. The latter is more often applied in library catalogues and article databases. Consequently, queries must be formulated differently in these two types of systems. It was assumed that full-text search would be easier for dyslexics, since such search systems allows for a wider range of query terms, and consequently the possibility to avoid potential difficult words by replacing them with synonyms. It was therefore expected that the effect of dyslexia may vary according to search system, with the largest impact in systems without full text search.

Cucerzan and Brill (2004) have reported that spelling errors are more frequent and severe in queries compared to word processing, and approximately 10-15% of the queries submitted in search systems contain errors. However, search systems vary in their tolerance level for misspellings. Some library catalogues and article databases have no tolerance for errors, and return empty result lists for misspelled queries. Other systems, such as most Web search engines, provide a result list of what is assumed the correctly spelled query. Consequently, the impact of dyslexia on search behaviour and search performance may be related to the system's ability to retrieve relevant results despite erroneous queries. This is the starting point for H_3. In a study of children and their use of online catalogues, Borgman, Hirsh and Walter (1995) concluded that searching was regarded as difficult due to high requirements for correct spelling. These findings may also apply to adult users with impaired writing skills such as dyslexics.

Some search user interfaces have implemented query-building aids to support users in query formulation (Hearst, 2009). For instance, search systems such as Google displays a red line beneath misspelled words in the search box. Another common query-building aid is the autocomplete, or query refinement function, which is based on word prediction and suggests terms while the users input queries. Consequently, the autocomplete function reduces the need for inputting complete words or long queries. An assumption for H_2 is that such functions are particularly helpful for dyslexics. The user can input the first characters of the query and subsequently select the appropriate query in the list of suggestions, thus reducing the need for correct spelling.

Although query refinement mechanisms can reduce cognitive load, it has been reported that users may be reluctant to utilise the autocomplete function (Dennis, McArthur & Bruza, 1998) because it is difficult to decide which of the suggested queries that may be most useful. Ward, Hahn and Feist (2012) concluded that autocomplete functions are mostly applied for spelling correction. Moreover, White and Marchionini (2007) have suggested that such functions are more useful in exploratory tasks and at the beginning of the search process.

It has also been suggested that autocomplete suggestions should be displayed in real-time, before the first results are retrieved, rather than in retrospect for users to notice them. However, White and Marchionini (2007) claim that real-times suggestions may also affect users negatively, and cause incorrect search paths.

When the user has formulated a query and submitted it to the search system, the documents in the result list must be assessed according to relevance. However, Belkin (2000) has pointed out that users do not always know which documents that are most useful. Moreover, the content of result lists is related to query formulation and the abilities of the system to retrieve the proper number of relevant documents. Consequently, if the user struggles with query formulation, the search process may continue with difficulties caused by irrelevant or confusing result lists.

The relationship between precision and recall is a challenge with query formulation. Precision refers to the number of relevant documents retrieved as a portion of the number of retrieved items (Buckland & Gey, 1994). In other words, precision refers to how many of the documents in a result list which are actually relevant. Recall is the number of retrieved relevant documents as a portion of all relevant items in the database (Buckland & Gey, 1994). Recall is thus a measure of how many of the relevant documents in the system's database which were actually included in the result list. However, there is a trade-off between precision and recall. Precision decreases with increasing amounts of documents in the result lists, which actually improves recall. In contrast, a decreasing number of items in the result lists improves the precision level but reduces recall. Short and quite general queries may typically result in long and imprecise result lists, while longer and more specific queries often entail shorter lists with higher precision.

According to Turetken and Sharda (2005), result lists may be overwhelming in size and content, and difficult to assess. Although result lists may contain thousands of Web pages, users rarely look at more than the top listed documents. However, since there is not always a correspondence between the relevance judgments of users and the systems' relevance ranking algorithms, Hariri (2011) has suggested that users should examine at least three or four pages with results.

Results are typically presented as textual lists, although other visual representations have been suggested, such as zoomable maps with clustered results (Turetken & Sharda, 2005). The inclusion of graphic content in result lists is addressed by H_1 and investigated in EXP-I, with the purpose of investigating whether the inclusion of graphic content may improve search performance.

2.4.2 Information retrieval and dyslexia

The introduction of a user-centred paradigm entailed an attention shift from systems to actual users. However, Hepworth (2007) has emphasised that these studies included little user diversity, with an overrepresentation of students or researchers, communities of common practise, people who share common needs and practitioners. Users with impairments were absent in the information seeking behaviour research up to the last few years. According to Beverly et al. (2007), disabilities are not included in any of the general user models.

Most existing studies on information retrieval concerning disabilities have addressed visual impairments, for instance Williamson and Schauder (2000), Craven and Brophy (2003), Beverly et al. (2007), Davies (2007), Chang and Chang (2010) and Sahib, Tombros and Stockman (2011). According to MacFarlane et al. (2010), cognitive disabilities have received less attention.

A few studies have addressed general cognitive abilities. For instance, Ford, Miller and Moss (2001) studied individual differences in online searching according to cognitive styles, prior experience, study approaches, Internet perception, age and gender. Cognitive style refers to the phenomenon where users apply different strategies when searching and processing information. Ford et al. (2001) found that retrieval effectiveness was related to, among others, low cognitive complexity levels.

Czaja, Sharit, Ownby, Roth and Nair (2001) investigated, among others, the effect of differences between cognitive abilities on task performance among users of different ages in a health insurance information database. They found that the older users were not able to compensate for reduced cognitive abilities in working memory. However, Czaja et al. (2001) suggested that a change of interface design may decrease the load on working memory.

Kim and Allen (2002) investigated if cognitive abilities, such as different levels of logical reasoning, spatial scanning and perceptual speed, could be used as predictors of students searching for online information. They found no significant effect for cognitive abilities within different levels of search activities and outcome, or in the interaction between search engines and cognitive abilities. However, Kim and Allen (2002) concluded that users with lower levels of cognitive abilities applied similar search strategies on all tasks compared to participants with higher level abilities who searched differently according to task.

Al-Maskari and Sanderson (2011) studied how user characteristics, such as cognitive skills and experience, affected search effectiveness. They reported that users with more experience found significantly more relevant documents than participants with fewer years of experience, but they used approximately the same time locating the first document. Further, there was a correlation between perceptual speed and search effectiveness, and users who scored high on perceptual speed tests used less time finding the first document compared to users who scored below the median of the tests. However, there were no differences between the two groups concerning number of documents retrieved. Al-Maskari and Sanderson (2011) reported no differences in the users' satisfaction regarding the results, although these varied according to search effectiveness.

Several general studies on dyslexia have included dyslexics and their relation to libraries. Riddick et al. (1997) interviewed dyslexic students who mentioned difficulties navigating in the physical library and understanding the reference system, a finding supported by Helland (2002). Helland also found that dyslexic students were afraid of going to the library due to the expected failures and embarrassment related to using the library services. This finding is coherent with Mortimore and Crozier (2006), who also reported that dyslexic students may have difficulties in using physical libraries, and thus require extra library support.

In a study by Habib et al. (2012) dyslexic students reported challenges in using search functions in the virtual learning environment applied at their university, because the system did not allow for spelling mistakes. The latter finding is in accordance with Al-Wabil et al. (2007). Rutledge (2002) investigated how the libraries in England and Wales accommodated dyslexic users. Rutledge found that although dyslexic library users may have different needs from other users, and information about dyslexia is commonly held in libraries, dyslexics are not included in policy documents. Rutledge (2002) concluded that dyslexics are excluded from public library facilities.

Of the few studies which investigate the effect of dyslexia on information seeking, MacFarlane et al. (2010) compared the search behaviour of five dyslexics and five controls in the test retrieval system Okapi. MacFarlane et al. (2010) found that users with dyslexia exhibited fewer search iterations, reviewed fewer documents and took more time on each search than the control group. No participants made spelling errors in this study.

MacFarlane et al. (2012) studied the search behaviour of eight dyslexics and eight controls with an emphasis on phonological short-term memory. Control users classified a higher number of documents as irrelevant compared to dyslexics. The authors also found a correlation between short-term memory capacity and the ratio of documents classified as irrelevant. MacFarlane et al (2012) reported that users with a higher phonological memory-capacity judged more documents as irrelevant compared to users with reduced memory-capacity. MacFarlane et al. (2012) also concluded that attention must be paid to the computer knowledge among participants, since computer skills were also found to impact the results.

2.4.3 Information seeking theories

Beverly et al. (2007) have reported a lack of information seeking theories and models considering impairments. However, some theories may have potential transfer value to dyslexics, although dyslexia is not specifically included in the original framework. The models included in this section are Wilson's model of information behaviour, Library anxiety, The principle of least effort and Information poverty. Gross' (2005) theory on imposed queries is also included, since this topic relates directly to the experimental design in EXP-II and EXP-III.

In **Wilson's** (1999) **model** (see Figure 6), information-seeking starts with the context of the information need. This is followed by activating mechanisms and intervening variables leading to different types of information behaviours. The model is circular, since information-seeking may result in a new or modified information need, which leads to a new round of information searching.

Several elements in Wilson's model are applicable in this study. First, the information-seeking behaviour includes active and ongoing search. Second, the intervening variables, described as potentially supportive or preventive (Wilson, 1999), include psychological and demographic characteristics. The cognitive profile of dyslexia, such as decoding errors, spelling difficulties and reduced short-term memory capacity (see Section 2.2.2), can be included in the psychological characteristics. Another relevant variable is source characteristics, which may include advantages or shortcomings in search systems, for instance the presence or absence of query-building aids and the tolerance level for spelling errors.

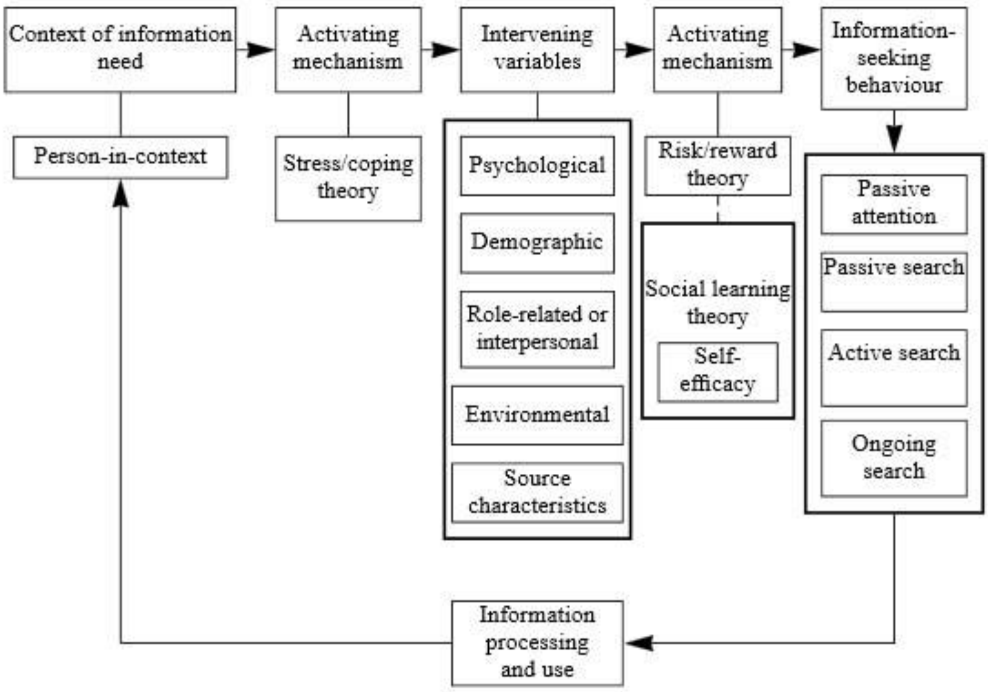


Figure 6: Wilson's (1999) model of information-seeking behaviour. Reproduced with permission by T. D. Wilson.

The concept **library anxiety** was introduced by Mellon, who found that students frequently discussed their feelings concerning the library when they were asked about difficulties related to information search (Katopol, 2005). Students often felt powerless in their encounters with the library, and mentioned feeling lost, afraid to approach the library staff and unable to navigate the library.

Mellon saw a need to identify the students' anxiety and find an approach to decrease these feelings. Consequently, she developed *The Library Anxiety Scale*, an attempt to quantify feelings towards the library based on five dimensions; barriers with staff, affective barriers, comfort with the library, knowledge of the library and mechanical barriers. This was later expanded upon in *The Multidimensional Library Anxiety Scale*, where for instance gender, online source preferences and use of off-campus information resources were added to the original scale (Katopol, 2005).

A fear of approaching the library is well documented among dyslexics (see Section 2.4.2). Consequently, library anxiety may be an appropriate description of many dyslexics' attitude towards information retrieval. The *library anxiety* is particularly interesting in the context of higher education. Most users have access to online information sources at home. Consequently, students can locate information without physically approaching the library. Persons with library anxiety may thus be particularly motivated to solve information needs unassisted. However, if these users frequently encounter inaccessible systems, information searching can become an insurmountable task.

Riddick, Sterling, Farmer and Morgan (1999) have reported that many users with dyslexia are anxious and self-conscious and struggle with low self-esteem issues. Consequently, accessible search systems and positive search experiences could improve the confidence among dyslexic users. Reducing library anxiety may therefore be significant for this user group and thus reduce the effort needed to locate academic and scholarly information.

In **The Principle of Least Effort** (PLE), also referred to as **Zipf's law**, the main principle is that users will solve tasks in a manner that requires the least amount of work (Case, 2005). This theory may be particularly applicable to dyslexics, who according to Caroll and Iles (2006) are assumed to develop good coping strategies. One common coping strategy is avoidance based coping (Alexander-Passe, 2006).

A study of dyslexic entrepreneurs by Logan (2009) reported better delegating skills among dyslexics than controls, and that dyslexics learned at an early age to delegate tasks they found difficult to complete. It is also claimed that both dyslexic adults and children generally avoid tasks which may reveal their dyslexia (Alexander-Passe, 2006; Logan, 2009). According to Riddick (1996), common strategies are to avoid words which are difficult to spell or to ask for help from others. Dyslexics are also found to avoid new learning situations due to little confidence in their own abilities (Riddick, 1996).

Although introduced in 1949, Zipf's law still seem valid, especially after the introduction of the Web. There is an overwhelming amount of online information, and users may not always choose the proper sources. For instance, according to Jamali and Asadi (2010), Google Scholar is the preferred information resource for many scholars, although searching several academic databases directly may be more proper. However, searching several databases separately would require more queries and longer search sessions. Zipf's law may also be seen in the number of users who are reported to rarely assess more than the top ranked documents in result lists (Granka, Joachims & Gay, 2004).

The concept of **information poverty** was introduced by Chatman (1996) in a framework called *the sociology of knowledge,* including the diametrical opposite concepts *information insiders* and *information outsiders*. Ethnographic approaches were applied while studying marginalized user groups, such as poor people, female prisoners, janitors and the elderly. Chatman (1996) found that there were barriers between the worlds of the insiders and the outsiders, and that outsiders were not expected to comprehend the world of the insiders

Chatman's work may at first not seem applicable to dyslexic users, since her work was primarily applied to different social classes. However, Hersberger (2005) has suggested that her theories may be adapted to other marginalized groups. If search systems are not accommodating the needs of dyslexic users, dyslexics may thus be considered information outsiders, and may in several contexts become a marginalized group (Macdonald, 2009).

Chatman introduced four key concepts: secrecy, deception, risk-taking and situational relevance. From these concepts six propositions were derived, of which two may be relevant concerning dyslexics: «*Information poverty is determined by self-protective behaviors which are used in response to normal norms*» and «*A decision to risk exposure about our true problems is often not taken due to a perception that negative consequences outweigh benefits*» (Chatman, 1996).

According to Hersberger (2005), Chatman's work has been cited by researchers working on information seekers who apply certain kinds of protective behaviour which may affect their access to useful information. Both propositions introduced above may be in accordance with dyslexic users who are reported to apply avoidance strategies to not reveal their dyslexia (Alexander-Passe, 2006; Logan, 2009). This assumption is supported by the small portion of online texts produced by dyslexics found by Baeza-Yates and Rello (2011). Further, the exposure risk may also be related to the findings that dyslexic users avoid using the library due to fear of failure, as reported by Helland (2002), thus reducing their means for retrieving essential information.

Gross (2005) developed a theory concerning **imposed queries**. Gross divided queries into two categories: self-generated and imposed queries. Self-generated queries originate in contexts surrounding the user. Imposed queries are not formulated by the user, but are either provided by someone else (for instance an assignment at school) or acquired on behalf of somebody (for example a parent searching for information for a child).

Gross (2005) stated that a successfully solved imposed query depended on several factors: the quality of question transfer, characteristics of the agent, imposer and intermediary, the relationship between the imposer and the agent and availability of resources. Gross' theory is especially applicable in EXP-II and EXP-III, where users are given predefined search tasks formulated by the experimenter, all of which may be classified as imposed queries. Characteristics of the agent include reading abilities, short-term memory capacity, information-searching skills and computer skills. The quality of information transfer may have implications for experimental design in the formulation and presentation of tasks, which will be discussed further in Chapter 3.

2.5 Visual search

Information retrieval is an activity which involves visual search, especially during result list assessment. H_1 addressed result list design with an aim to explore whether graphic content may assist dyslexic users when navigating result lists, and was investigated through a visual search experiment.

2.5.1 Visual search strategies

Visual search is the activity of scanning a visual scene for a target with particular features, and has been applied as a methodology for studying both visual and attentional processes. According to Findlay and Gilchrist (2003), most studies follow the traditional visual search paradigm. A participant is typically presented with a display containing a target item and several distractor items, asked to search for a target and indicate whether the object is present or not. Typical performance measures are reaction time and error rates.

A traditional search task may for instance be to locate a B in a display consisting of the letter P (see Figure 7). This is in contrast to the stimulus applied in EXP-I, which was more similar to actual user interfaces consisting of icons, words (see Figure 7) or both content types in dual-modality displays. Although the tasks in EXP-I are more experimental than actual result lists, they were assumed to more closely imitate an actual user interface compared to traditional visual search tasks.

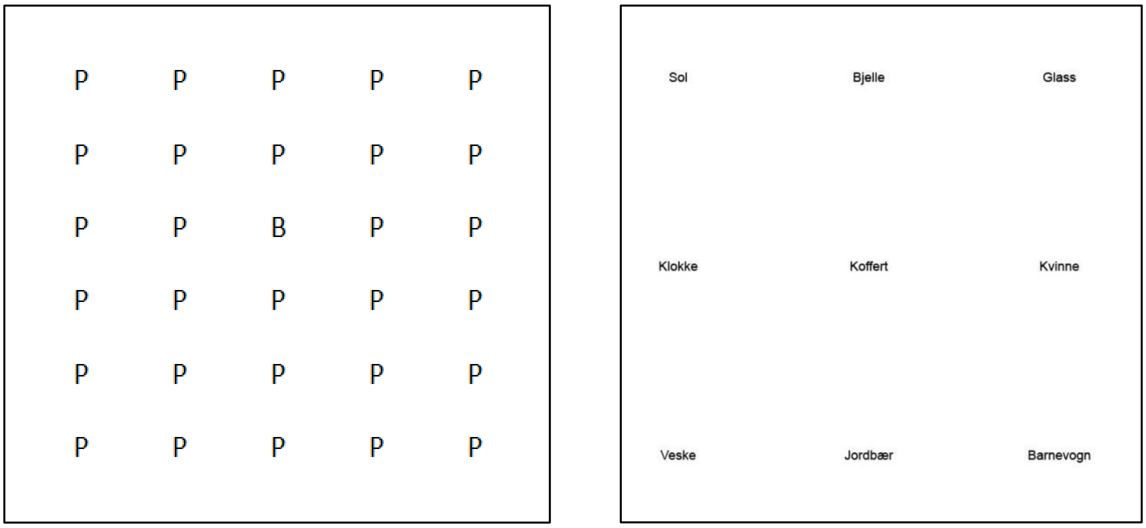


Figure 7: Traditional visual search task (left) and excerpt from EXP-I WORD-task (right)

In visual search, targets are identified either by one separate feature or by a conjunction of features. Treisman and Gelade (1980) introduced a feature-integration theory (FIT) stating that visual search may follow two different search paths: parallel search and serial search. If there is one separable feature, for instance colour or orientation, the target will have a pop-out effect. In such cases several items can be processed at a time, thus resulting in a parallel search or feature search. The task to the left in Figure 7 is an example where such a search strategy may be successfully applied. If more than one feature is needed to identify a target, attention must be directed at each item. This strategy is referred to by Treisman and Gelade (1980) as a serial search or conjunction search.

A variation of the classical visual search paradigm is word search, where a participant is asked to locate a target word in a set of distractor words. Ojanpää, Näsänen and Kojo (2002) claim that words represent a relatively complex stimulus, and have therefore suggested that such tasks require serial search. Wolfe, Cave and Franzel (1989) suggested an alternative model to the FIT-theory, referred to as guided search. According to Wolfe et al. (1989), serial search may be guided by parallel processes, which can be used to divide the objects into possible candidates for targets and distractors. Attention will first be directed to the area where the target is most likely to be located. If the target is not found, attention will be directed to the next potential candidate.

2.5.2 Eye movements

Eye movements are widely used to investigate cognitive processes during visual search (Rayner, 2009). This is in contrast to previous decades, where eye movements were not assumed to be of particular importance in such studies. According to Findlay (2004) it is well documented that most search tasks involve multiple fixations. When the human eye is directed towards a visual scene, retinal images are aligned in the central area of the fovea (Wright & Ward, 2008).

The centre of the fovea provides the most detailed visual input, while visual abilities decline in the outer areas. Consequently, to process objects in detail, eye movements are necessary, so that the target is within the centre of the fovea. The region in which discrimination of targets can be made is also referred to as the conspicuity area, the visual lobe, and visual span in reading (Findlay & Gilchrist, 2003).

Eye movements in a visual search task may depend on the type of task. According to the FIT-theory introduced by Treisman and Gelade (1980), tasks may be solved either by parallel search or serial search. In the context of eye movements, a target may be located by a single fixation in parallel search, while serial search requires several fixations and saccades, especially if there is a large display size (Findlay & Gilchrist, 2003).

Bertera and Rayner (2000) have reported that the size of the visual lobe varies according to, among others, target type, visual acuity and task difficulty. Irwin (2004) has pointed out that the visual span may also be affected by characteristics such as age. Ojanpää et al. (2002) claim that the horizontal word identification span is approximately ten character spaces, which corresponds to 1-2 words, while the vertical span may enable the identification of 4-5 words. However, the word identification span may be affected by factors such as letter size and word spacing.

If a target is surrounded by other targets, visual perception becomes more difficult, a phenomenon which is referred to as lateral masking (Findlay, 2004). The capacity of visual processing is also limited. Consequently, there is a cognitive selection where the objects regarded as most relevant are processed, while distracting objects are ignored. Ward (2001) refers to this phenomenon as selective processing.

Another phenomenon affecting object recognition is crowding. Object recognition consists of two phases: feature detection and combining these features into objects. According to Pelli and Tillman (2008), if objects are arranged too close together, features from several objects may be combined erroneously, thus causing crowding and failure in object recognition. Vlaskamp and Hooge (2006) studied the effect of crowding on eye movements. They found that increased crowding led to longer fixation durations, a higher number of fixations and decreased saccadic amplitude. The search performance also became worse as the crowding increased. Moores et al. (2011) suggested that users with dyslexia are assumed to be more affected by crowding than non-dyslexics.

Peripheral vision is also assumed to be involved in visual search. However, Findlay (2004) claims that peripheral vision can merely be used to detect possible targets in displays with multiple isolated items. Scanning a display looking for a target often involves short fixations, and elements may be scanned superficially, just to eliminate the object as the target the person is looking for. This has been referred to as intermediate stages of perception (Neisser, 1964).

Attention is another much-debated topic in visual search. According to Findlay and Gilchrist (2003), researchers agree upon two types of attention: overt and covert attention. However, it is not agreed upon how these types of attention interact in visual search. Overt attention involves an eye movement to a specific target which a participant wants to pay attention to, while covert attention is an attention shift towards a target in the visual periphery without any eye movements (Findlay & Gilchrist, 2003). Consequently, covert attention is not observable in eye data.

Covert attention may have implications for experimental design on visual search tasks. To avoid parallel search, which will not be directly observable, items should be displayed in relatively small sizes with enough object spacing so that only one object can be examined per fixation (Peterson, Kramer, Wang, Irwin & McCarley, 2001). This is discussed in more detail in Section 3.4.1.

It is disputed whether an eye movement always is related to a purposeful attention shift. Henderson (2012) claims that attention is typically tied to fixations. Wright and Ward (2008) have reported that most researchers using eye tracking assume that there is a connection between attention and where the eye is fixated, although it is acknowledged that this may not always be so.

According to Holmqvist et al. (2011), the eye-mind hypothesis developed by Just and Carpenter (1980) is commonly applied in eye tracking studies. This hypothesis states that there is a direct relationship between what is being fixated on and what is cognitively processed. Regarding words, Just and Carpenter (1980) present two assumptions: the immediacy assumption and the eye-mind assumption. The first refers to the idea that the reader tries to interpret the word immediately when it is encountered. The second assumption states that the eye rests on the word in the same timespan as the word is processed.

Regarding eye movements, both fixations and durations have been related to search times and search performance. For instance, it has been suggested that fixation durations (or dwell times) may reflect cognitive load or task difficulty. Ojanpää et al. (2002) found that average fixation lengths depended on the number of items within a span. Benjamins, Hooge, van Elst, Wertheim and Verstraten (2009) found that search times increased with the number of distractor items in a display. According to Becker (2011), dwell times may also be affected by the similarities between the target and the distractors, and the target size. Phillips and Edelman (2008) found that improved search performance was more closely related to saccadic metrics than fixation duration. They found that search speed mainly increased when participants processed information from a larger area of the visual speed instead of processing information from a certain area more quickly.

Nothdurft (1999) claims that in tasks where a pop out effect occurs, search times are independent of the number of objects, while search times increase according to set size in displays with more similar targets. It has been assumed that unfamiliar objects require more effort to evaluate, and will therefore result in longer search times. However, if both the target and the distractors are either familiar or unfamiliar, Greene and Rayner (2001) reported that search rates are slow.

Studies have also addressed whether eye movements in word decoding may differ in reading and single word identification. According to Vitu, McConkie, Kerr and O'Reagan (2001), previous research has found that words are most quickly identified when the fixation is in the middle of the word, also referred to as the Optimal Viewing Position effect. Such eye movements reduce the need for additional fixations to decode a word. However, Vitu et al. (2001) studied if this effect was also applicable for reading, and found that fixations were longer in the middle of words than towards the beginning or end. They also suggested that there is a correlation between fixation durations and the frequency of the word in the general language corpus.

2.5.3 Dual-modality displays

Brante et al. (2013) have summarised conflicting results in previous research on the usefulness of visual support for dyslexic users. The visual search experiment in this study applied dual-modality displays containing icons and words.

Icons are pictographic representations applied in several contexts, for instance on signs, product labels or in computer user interfaces. Huang, Shieh and Chi (2002) have summarised several advantages with icons: they are memorable, can be easily interpreted and they may be more universal than text, since language is not involved. The interpretation of icons can be influenced by user characteristics such as cultural background, gender and age (Berget & Sandnes, 2015b). Consequently, icons must be universally understandable and identifiable for a diversity of users. Huang et al. (2002) propose five important factors for usable computer icon design: styling quality, message quality, meaningfulness, familiarity and the icon should work well as metaphor.

Niemelä and Saarinen (2000) have reported a general user preference for icons. Nielsen (1990) found that users generally preferred icons over miniatures, although there were no significant differences in user performance. Kacmar and Carey (1991) investigated the application of menus including text, icons and dual-modality displays. They found no differences in time usage, but that menus with both content types resulted in the least incorrect target choices. Niemelä and Saarinen (2000) reported improved search times when searching icons compared to words, but search times depended on set-size and layout, namely whether icons were grouped or randomly positioned. This is in contrast to Benbasat and Todd (1993) who concluded that there were no advantages with icons compared to text-based representations.

Lindberg and Näsänen (2003) investigated how icon size and spacing affected visual search times. They found that search times increased with set-sizes, but that icon spacing did not affect search performance. However, when icons became smaller, search times increased significantly. Lindberg and Näsänen (2003) concluded that icons should always be displayed some distance apart, with a preferable inter-element spacing of one icon.

According to Dampuré, Ros, Rouet and Vibert (2014), most visual search research has included shapes and symbols. However, a few studies have focused on verbal material, and found that if target words are surrounded by orthographic distractor words, search times increase. This was confirmed by Dampuré et al. (2014) in an eye tracking study on how people search for single target words.

According to Huang (2012) and Huang, Bias and Schnyer (2014), icons are processed differently from text, and require more neurological resources. The studies summarised in the previous sections have reported contradicting results concerning the usefulness of icons. However, user performance and user preferences do not always correlate. Users have been documented to not always apply the most efficient strategies, and Nielsen (1990) claims that users do not always prefer the solution in which they perform the best. This is in accordance with for instance Williams and Hennig (2015b), who found that users with learning disabilities preferred designs with large text and images, although these conditions did not facilitate the fastest search times.

The icon studies referred to above have investigated users in general, without an explicit focus on users with reading difficulties. However, other studies have directed attention particularly to dyslexics and their use of dual-modality displays. According to Santana et al. (2012), dyslexics tend to place more emphasis on images than words. This is supported by Houts, Doak, Doak and Loscalzo (2006), who found that pictures can increase attention, recall and comprehension if they are closely linked to text. Houts et al. (2006) emphasise the importance of minimizing distracting details in pictures and that simple language should be used in connection with pictures. Further, the design process should include people from the intended audience, and there should be an evaluation of the material with and without the inclusion of pictures.

Other researchers have reported contradicting results, and conclude that pictures may distract attention away from the text and negatively affect reading comprehension. Beacham and Alty (2006) studied the learning outcome of students in different media types and found a contrast between what the dyslexic students preferred and how they performed. The majority of the dyslexic students claimed that a combination of text and diagrams were easiest to follow. However, educational material including text-only provided most improved learning (Beacham & Alty, 2006). This finding is supported by Brante et al. (2013), who studied the use of learning materials containing text-only or text and picture materials, and concluded that integrated pictures in learning materials did not help dyslexic students in their comprehension of the material.

Williams and Hennig (2015a) investigated the most efficient menu orientation for users with learning disabilities. Participants were asked to locate frequent English words in lists laid out horizontally or vertically with an illustrative graphic. Williams and Hennig (2015a) analysed search times in tasks with different conditions and found that horizontal menus were most efficient. In another study, Williams and Hennig (2015b) explored how to optimize Web sites for users with learning disabilities to reduce navigation difficulties. They found that users performed the best in horizontal menus and conditions with small-sized text. Further, the use of pictorial representations was preferred by the users, although they did not seem to aid comprehension, among others, due to difficulties finding appropriate images to abstract concepts.

2.5.4 Visual search and dyslexia

Several studies have addressed visual search skills among dyslexic users. However, researchers do not agree upon whether dyslexics are impaired in all visual search tasks, certain task types or if search performance varies according to subtype of dyslexia.

Skottun and Skoyles (2007) claim that dyslexic users exhibit sensory deficits. According to Prado et al. (2007), abnormal eye movements while reading sentences, single words and pseudo-words are well documented. People are assumed to vary their abilities in making use of information in the centre and periphery of the visual field, and some dyslexics are reported to have a preference for the peripheral visual field (Schneps, Rose & Fischer, 2007). Although this preference may cause visual deficits, Schneps et al. (2007) suggested that these preferences may enhance visual comparison abilities. Facoetti, Paganoni and Lorusso (2000) found that dyslexics distribute their attention diffusely, and may therefore have difficulties narrowing their attention. Consequently, dyslexics adopt a more distributed strategy in visual search tasks.

Prado et al. (2007) studied French dyslexics who applied a higher number of rightward and leftward fixations and longer fixations during reading than controls. However, these differences in fixations did not impact the search performance in visual search. Prado et al. (2007) also found that controls displayed more rightward fixations in visual search compared to reading, which indicated a larger attention span during reading than visual search. Huang et al. (2008) studied the eye movements of Chinese dyslexics, and reported that although dyslexic children performed equally well as controls on a picture vocabulary test, they performed worse in picture searching. Wright, Conlon and Dyck (2012) compared visual search performance among 70 dyslexic children and 52 controls, and discovered slower serial search times in the dyslexia group.

Visual search performance among dyslexics is assumed to be affected by several factors. For instance, Moores et al. (2011) identified stimuli characteristics, such as small spacing and increased set-sizes to affect the dyslexic users, and concluded that dyslexics exhibit large effects of crowding. The set-size effect was partly confirmed by Vidyasagar and Pammer (1999), who found no significant differences in mean search times between dyslexics and controls with set-sizes of 10, 24 or 36 items, but a significant difference with 70 items.

Sireteanu et al. (2008) concluded that dyslexic children were impaired in serial visual search, but there were less differences compared to controls on feature tasks, among others, in letter-like stimulus. Sireteanu et al. (2008) also found that the children with dyslexia were more easily fatigued than controls. Although Sireteanu et al. (2008) reported improvements with age, Boer-Schellekens and Vroomen (2012) have described impairments in visual search tasks in young adults and adults. Jones et al. (2007) studied adults with dyslexia, and found more errors in the dyslexia group compared to controls on both visual search tasks and symbol tasks.

Visual search has also been studied in the context of different types of dyslexia. For instance, Iles, Walsh and Richardson (2000) concluded that dyslexic users with motion coherence deficits were impaired in serial search tasks, but not on parallel tasks. However, dyslexics with no motion coherence impairments were unimpaired in visual search.

2.6 Summary

The research literature discussed in this chapter confirms the gap in existing research on cognitive disabilities in the context of universal design. Cognitive impairments have also received less attention in accessibility guidelines such as WCAG (W3C, 2008), which according to Santana et al., (2012) may be an obstacle in developing a universally designed Web. Studies have suggested that the Web in general and search systems in particular may be inaccessible for dyslexic students (Al-Wabil et al., 2007; Habib et al., 2012; Santana et al., 2012). The purpose of this study is thus to acquire detailed empirical knowledge about how search system design may better accommodate dyslexics during query formulation and result list assessment. This knowledge may be applied to enhance the accessibility of both Web search user interfaces and library catalogues.

Accessible libraries and universal design have been discussed for the field of library- and information science, with a focus on accessible Web sites. However, there seems to be a lack of studies on accessible library catalogues and search systems. There is also a gap in the research literature on information searching behaviour studies that address disabilities in general, and dyslexia in particular. One aim of this study is reduce this gap.

The dyslexia research discourse is extensive, and there seems to be consensus on the cognitive profile of dyslexia (Snowling, 2000). Mortimore and Crozier (2006) argue that a weakness in the dyslexia research is that most of the dyslexia research has focused on children, while students in higher education have been less addressed. Consequently, there may be a gap in the research on adults with dyslexia, which is the participant group addressed in this study, described in more detail in Chapter 3.

Chapter 3: Methodology

Chapter 2 introduced several issues affecting experimental design, such as the cognitive profile of dyslexia and recruiting difficulties associated with impaired users. Moreover, research on eye movements and attention addressed the necessity of appropriate stimulus design in visual search experiments, regarding layout, object spacing and set sizes. This topic is particularly relevant for the visual search experiment (EXP-I). Search user interfaces functionalities were also discussed, and have implications for the choice of search systems and tasks in the information retrieval experiments (EXP-II and EXP-III). This chapter concerns methodology and validity.

Section 3.1 discusses the methodologies applied, and includes a short introduction to eye tracking, apparatus and ethics. Section 3.2 describes participants, recruitment procedures, inclusion criteria and participant characteristics. Next, Section 3.3 introduces five screening tests, namely a visual acuity test, a dyslexia screening test and three cognitive tests for short-term memory and attention. The experimental design is detailed in Section 3.4.

3.1. Research methods

Methodology concerns the choice of methods and the forms in which the methods are applied, involving techniques and procedures for research. Methods are commonly divided into quantitative or qualitative approaches. According to Crotty (1998), the main distinction between these approaches occurs at the level of methods, not at an epistemological level or within theoretical perspective. However, quantitative and qualitative methodologies are frequently described as polar opposites. Crotty (1998) argues that this division is not accurate, since much qualitative research has been conducted in an empirical, positivistic manner, while quantification of data is not only limited to positivistic research.

Beins (2009) describes the scientific method as «*the mode of accepting knowledge based on empirically derived data*». Most textbooks on research methods emphasise that scientific research should be objective, data driven, verifiable, replicable and valid. In addition, Beins (2009) argues that research should be public and useful. Usefulness has been an important motivation behind this study. An overall aim was to investigate the information retrieval of dyslexics and to get empirical knowledge on how to make accessible search user interfaces for this user group. The purpose is to communicate these findings to decision makers, developers and designers.

Three hypotheses were explored: whether the inclusion of graphic content in result lists may enhance the search process for users with dyslexia (H_1), if dyslexics rely on query-building aids during query-formulation (H_2) and whether the impact of dyslexia on information search behaviour varies with the tolerance level of the system (H_3). Several methods could address these questions, such as case studies, interviews, questionnaires, observation, search logs, case studies, heuristic evaluation, screen recording and eye tracking.

Methods relying on self-reporting, such as interviews, questionnaires and search logs could provide input on search preferences and certain aspects of search behaviour without affecting the participants regarding the choice of systems or the tasks. However, users may not be aware of certain behaviour or not remember shortcomings in user interfaces when asked. In addition, it has been reported that users do not always have a preference for the solutions where they actually perform the best (Nielsen, 1990; Williams & Hennig, 2015b). Moreover, asking users to fill out questionnaires or write search logs may also contribute with a substantial workload for dyslexic participants.

Other options were case studies and observation, which did not rely on self-reporting, and that allowed for observation of users in natural contexts. Such methods would also remove the potential challenges associated with imposed queries (Gross, 2005). However, this method is quite time-consuming, and would have significantly reduced the sample size. It would therefore have been difficult to get a large enough sample size to make any generalisations. Recruitment may also be more difficult, since participants might be less reluctant to be followed closely by a researcher over time than to attend a certain number of scheduled sessions in a lab.

A heuristic evaluation of search systems would have been an alternative. However, W3C (2008) states in the introduction to WCAG 2.0 that these guidelines do not fully accommodate the needs of users with cognitive impairments. This is in accordance with Rømen and Svanæs' (2012) findings. Consequently, the heuristics applied in an evaluation may not be sufficient to reveal all shortcomings. Further, the experimenter conducting the evaluation may not have sufficient knowledge or experience with dyslexia to identify all potential difficulties. Consequently, a heuristic evaluation of search user interfaces would probably not have been adequate to identify all relevant design issues. Further, a heuristic evaluation would not provide answers to the hypotheses which were more directly addressing behaviour.

A more traditional experiment in a lab involving actual users without self-reporting was regarded as the most sufficient method to answer the three hypotheses. This decision was based on the assumption that dyslexic users are most suitable to identify and express their own needs. Further, according to Petrie and Bevan (2009), observed behaviour is a good indication of usability and accessibility, and enables the documentation of unexpected behaviour.

Eye tracking and screen recording were chosen as the main methods for data collection. Eye tracking is a frequently used method, and is reported to be quite objective (Poole & Ball, 2006). The experimental software for the eye tracker is also easily combined with keyboard-logging, and could therefore reproduce search logs that may be quantified and analysed statistically. A qualitative approach was applied through interviews before the experiments and during debriefing, to obtain data on variables such as previous search experiences and attitudes towards search systems (see Section 3.2.3).

The validity of a study may be affected by factors such as inclusion criteria and recruitment, experimental design and the methods for data collecting and analysis. Researchers frequently engage in methodological debates, and preferences may vary according to contexts and research questions. However, in some cases the researcher cannot always choose between all available methods due to the nature of the research question. For instance, randomisation of participants into groups is commonly applied in experiments, where users are non-systematically placed in different groups.

The aim is to have equivalent groups at the start of an experiment, to ensure internal validity (Beins, 2009) and be able to observe the effect of for instance a treatment. However, in other studies random assignment is not possible or applicable. Such experiments are often referred to as natural experiments (Shadish et al., 2002), and commonly apply a static-group comparison design (Beins, 2009). For instance, in a study with the aim of observing the effect of dyslexia, participants must be assigned into groups based on the presence or absence of a dyslexia diagnosis, and randomisation is therefore not applied in such contexts.

3.1.1 Eye tracking

Eye tracking is a method for measuring a person's eye movements, which according to Poole and Ball (2006) has become a common method for usability testing in HCI. Eye tracking has been used in several Web accessibility studies and in previous research on information retrieval and visual search. For instance, Pan et al. (2004) studied Web page viewing behaviour, while Byrne, Anderson, Douglass and Matessa (1999) addressed visual search of click-down menus. Buscher, Cutrell and Morris (2009) investigated where users allocate attention on Web pages, to predict the most salient regions of a web page. Goldberg, Stimson, Lewenstein, Scott and Wichansky (2002) studied design features for a Web portal application, while Cutrell and Guan (2007) addressed information usage in Web search. Beymer, Russell and Orton (2008) investigated the effect of font types and sizes on online reading.

Eye tracking has also been applied to study information retrieval. For instance, Granka et al. (2004), Aula, Majaranta and Räihä (2005), Rele and Duchowski (2005) and Guan and Cutrell (2007) applied eye tracking to study result list evaluation in Web search engines. Lorigo et al. (2008) addressed online search in three experiments, and concluded that eye tracking is a useful method for investigating user behaviour. Eye tracking has also been applied by MacFarlane et al. (2010) to study the effect of dyslexia on information retrieval.

Poole and Ball (2006) describe eye tracking as an important, objective technique, which allows for more in-depth analysis of the usability of user interfaces. Bruneau, Sasse and McCarthy (2002) have also addressed the use of eye tracking in HCI research. They concluded that this method may provide objective data on how visual design affects human performance. This view is supported by Jacob and Karn (2003), who describe eye tracking for evaluating user interfaces, among others because eye tracking technology has become much more reliable and accessible over the last years. Another advantage is that eye tracking produce quantitative measures, which may be analysed statistically.

Several types of eye trackers are available, and they facilitate different types of studies. For instance, a remote eye tracker can be used when a person is sitting in front of a computer, while eye tracker glasses allow participants to move freely.

Most eye tracking technologies applied today include software for estimating the point of regard based on the relationship between the position of the pupil and the corneal reflection (Poole & Ball, 2006). An infrared camera is directed towards the participant's face. The light that hits the retina is reflected back as a disk (see Figure 8). Poole and Ball (2006) refers to this phenomenon as the bright pupil effect. In addition, a reflex will appear when the infrared light hits the cornea (see Figure 8), which is the transparent layer that covers the eye. This is usually referred to as corneal reflection, 1st Purkinje reflection or glint (Holmqvist et al, 2011).

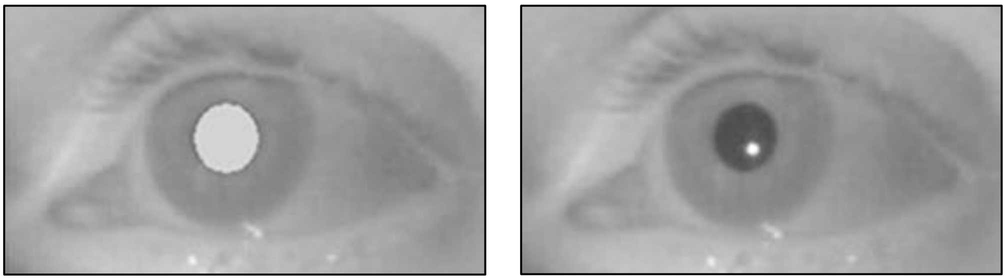


Figure 8: Bright pupil (left) and corneal reflection (right). Image by Gneo, Schmid, Conforto and D'Alessio, released in Creative Commons 2.0. Retrived from https://commons.wikimedia.org/w/index.php?curid=38802822

When initiating an eye tracking session, each participant completes a calibration procedure to ensure sufficient tracking quality. A common method is to display dots in different areas of the screen, one at a time. The participant is asked to fixate on each point until it disappears. According to Holmqvist et al. (2011), the accuracy values for the calibration procedure should be a viewing angle of approximately 0.2°, visual degrees, with a maximum deviation of 0.5° visual degrees

Holmqvist et al. (2011) report that eye tracking quality is affected by factors related to the environment and participants. Environmental factors include lighting conditions, noisy environments and the eye tracking technology, such as tracking speed and whether it records monocularly (one eye) or binocularly (two eyes). Holmqvist et al. (2011) also mention participant variables, such as the use of mascara, droopy eyelids and downward eyelashes, shadows and reflections in glasses, the use of contact lenses, wet eyes due to tears or allergic reactions or hair covering parts of the eyes.

3.1.2 Apparatus

The experiments were conducted in an interaction laboratory at Oslo and Akershus University College of Applied Sciences. Eye movements were recorded binocularly by an SMI remote eye tracker mounted below the screen at 250 hz, which equals a recording ratio of 250 measurements per second. An automated, low velocity five-point calibration was applied for all participants. The calibration displayed red dots on a white background to ensure similar luminance on the calibration and stimulus, to avoid changes in pupil sizes that could affect eye data quality.

The validation option was also activated, so that four additional points were added to the calibration procedure. A threshold was set at 0.5° visual degrees for all participants. No experiments were started until the calibration was valid. The lab was equipped with items such as eye make-up remover, eye lash curler, cleaning equipment for glasses, eye drops, hair elastics and hairclips to eliminate factors that could negatively affect eye data quality, discussed in Section 3.1.1.

The situational variables were controlled as much as possible, to ensure equal testing conditions for all participants. For instance, markers were placed on the table to ensure that the placement of the computer screen was the same for all participants and sessions. A chin rest was used to reduce head motion, and ensured a fixed viewing distance of 70 cm. From this distance, 1° of vision equalled approximately 43 pixels. A steady chair, without wheels, was used so the participants would not move significantly during the experiment. The equipment was arranged so that no light sources would affect the eye tracking quality. The same lightning source was used for all experiments, and curtains were closed to avoid sunlight.

A 21' flat screen Dell LED monitor was used to display full-screen stimuli, with the screen resolution set to 1680 x 1050 pixels. SMI Experiment Center 3.2.11, a visual presentation software, was used for constructing and running the experiments. The screen recording option and mouse/keyboard input was activated during all experiments. Search tasks in the information retrieval experiments (EXP-II and EXP-III) were solved in the browser Windows Internet Explorer version 9, with all functions related to prior use such as browsing history and cookies disabled.

3.1.3 Materials

Visual stimuli for EXP-I and instructions for EXP-II and EXP-III were created in Adobe Photoshop CS4. Verbal instructions were pre-recorded with the speech synthesis programme Voxit speak-app, with the voice option Kari. Pre-recorded instructions ensured the exact same instructions for all users. For the cognitive tests, the experimenter had notes for instructions and followed standardised procedures, to ensure that the tests were conducted equally for all participants.

Verbal instructions were used to avoid misinterpretations caused by decoding errors, and also prevented participants from seeing the spelling of target words or query terms. Verbal instructions may place a more cognitive load on the participants because they have to remember the tasks. This could be a challenge for dyslexics, who are often associated with reduced short-term memory capacity (Smith-Spark & Fisk, 2007). However, the task of reading the instructions was considered more cognitively demanding than memorising short instructions. The assumed discrepancy between reading and listening abilities among individuals with reading disabilities is also supported by the literature (Badian, 1999).

3.1.4 Ethics

The PhD project as a whole formally adheres to the ethical standards of the Norwegian Social Science Data Services (NSD), the Data Protection Official for Research for all Norwegian universities and university colleges (see Appendix G). NSD approved the overall study (project number 29348), including the experiments and screening tests, the recruiting letter, consent form, interview guide and methods for collecting and storing data.

The participants signed consent forms and were informed that they could leave the study at any time, without giving the experimenter any particular reasons for the withdrawal. They were also instructed to only answer questions they were comfortable with. Further, they were not asked to provide any sensitive documents concerning dyslexia diagnoses, but rather screened with a test in the lab (see Section 3.3.2).

All participants were allocated participant codes. A document with the mapping of participant numbers to personal contact information was the link between participants and the codes, and only the experimenter had access to this file. All digitalised data was stored on an encrypted disk. Physical sensitive data such as consent forms were locked in a safe, and only the experimenter had the access code. All the data was anonymised during analysis and publication of results.

It can be challenging to keep the anonymity of participants who are recruited from a relatively small population, such as dyslexic students. Some statements from the dyslexic users were therefore omitted from the papers, although they provided interesting information. This was done in cases where the content was regarded to potentially reduce the anonymity of the participant. The experiments in the lab were completed with a significant gap of time between each session, to prevent participants from meeting each other when leaving or arriving at the lab. The participants had no relations to the experimenter or any of the supervisors.

3.2 Participants

A set of inclusion criteria was applied for the selection of participants. First, the participants had to be students attending higher education programmes. Second, Norwegian must be the participants' first language, since most tasks were related to reading and writing, and language barriers could affect the results. Third, it was important that participants had adequate sight, so that reduced vision would not affect the results.

Visual acuity refers to the size of the smallest letter or symbol a person can see on a test chart (Rabbetts, 2007). The inclusion criteria were set to visual acuities of at least 0.6 on each eye separately and 0.8 with both eyes open, which according to Zhang, Bobier, Thompson and Hess (2011) are within the range regarded as normal vision (see Section 3.3.1). The final criterion was that dyslexics had to be formally diagnosed, while controls had to achieve a certain score on a dyslexia screening test (see Section 3.3.2).

Students from library- and information science were excluded due to their extensive training in information retrieval, which could influence the results of the information retrieval experiments. The control group was assembled to match the dyslexia group according to gender, age, field of study and year of study.

Dyslexic students may be regarded as high-functioning dyslexics, which may be a possible limitation of the study. However, this inclusion criterion was applied for several reasons. Students were assumed to have similar frequency of use and experiences with search. It was also necessary to include users sharing similar experiences with academic library catalogues, since such a system was used in the third experiment and is not commonly used outside the academic realm. Further, assembling an applicable control group became more manageable, since participants were selected from a relatively limited part of the population. Nevertheless, if users with more severe dyslexia had been included, the results may have come out differently. However, including only high-functioning adults with dyslexia reduced the risk of including people with additional severe learning disorders.

3.2.1 Sampling

According to Sutton et al. (2003), recruiting participants from vulnerable user groups may be difficult. Information concerning a dyslexia diagnosis is highly sensitive. Contact information for dyslexic students is thus not available to researchers, which is an important privacy protection issue. Consequently, the study relied on self-selected samples, where students volunteered to participate. Further, completing a higher education may be demanding for dyslexics (Helland, 2002), who may neither have the time nor the energy to participate in outside-curriculum activities. Moreover, a lack of self-esteem and fear of failure which are commonly reported for dyslexics (Caroll & Iles, 2006) may cause discomfort towards testing situations and affect the motivation for participation.

Participants were recruited through requests in classrooms, announcements in social media and on relevant Web pages. In addition, dyslexics were recruited through enquiries sent by Dyslexia Norway on behalf of the experimenter to students registered in their list of members. The office for students with special needs at one University and one University College were also asked to mention the study for potential dyslexic candidates. Participants for the control group were recruited through announcements in classrooms, social media and Web pages. All participants received a gift card valued 400 NOK after completing the experiments.

3.2.2 Participant variables

A total of 45 students attended the first lab session; 22 dyslexics and 23 controls. However, three individuals were excluded; two dyslexics due to reduced visual acuity and one matching control. The final sample size was 42, consisting of 21 students in a dyslexia group and 21 non-dyslexics in a comparison group. One of the dyslexics had to withdraw from the study after the first experiment. Consequently, the matching control was removed from the dataset in the second and third experiment, thus reducing the sample size to 40.

Goldberg and Wichansky (2002) reported that the typical number of participants included in eye tracking studies vary from 6 to 30 people. For instance, Granka et al. (2004) recorded the eye movements of 26 people performing 10 tasks, MacFarlane et al. (2010) had sample sizes of ten (five dyslexics and five controls), MacFarlane et al. (2012) studied 16 participants (8 dyslexics and 8 controls) and Rello, Kanvinde and Baeza-Yates (2012) included 22 dyslexic users. Consequently, the sample size should be comparable with similar studies.

The comparison group was matched against the dyslexia group according to gender, age, field of study and year of study (see Table 2). The participants represented different fields of study, for instance nursing, engineering and various subjects in humanities and social sciences.

	EXP-I (N=42)		**EXP-II & EXP-III (N=40)**	
	D-group	**C-group**	**D-group**	**C-group**
Female	57.1%	57.1%	60.0%	60.0%
Male	42.9%	42.9%	40.0%	40.0%
Age	24.0	23.4	24.2	23.6
Year of study*	2.3	2.3	2.4	2.4

Table 2: Participant variables, mean values

*Bachelor students are valued as year 1-3, master as 4-5.

The gender distribution of the sample is in accordance with the general student population, which in 2015 comprised 39.7% males and 60.3% females (NSD, 2016). However, Hawke et al. (2009) have reported a higher prevalence of dyslexia among men than women. Consequently, the male portion of participants could ideally have been higher. However, it was difficult to influence the gender distribution when relying on volunteering students. Moreover, it was hard to recruit volunteers who were comfortable with their dyslexia diagnosis.

Dual diagnoses are commonly found in the dyslexia population, as discussed in Section 2.2.1. Three of the dyslexic students and one control were diagnosed with ADHD or ADD, and one dyslexic had dyscalculia. Consequently, there were prevalence rates of 14.3% with ADHD/ADD and 4.8% for dyscalculia. These numbers are quite representative compared to the general dyslexia population, where these diagnoses are reported to appear in 18-20% for ADHD/ADD (Germanò et al., 2010) and 4-7% for dyscalculia (Landerl et al., 2009). Dyslexics with dual-diagnoses do therefore not seem to be overrepresented in the sample.

3.2.3 Pre-interviews

Pre-interviews were conducted before the information retrieval experiments. Participants were asked about topics such as formal training and search experiences and their attitudes towards the search systems applied in the experiments.

According to Kuhlthau (1991), formal training and prior experiences may affect search behaviour. Only one control user had received training in Web search engines. A higher number of students had been formally trained in bibliographic databases as a part of their study programmes (eight dyslexics and nine controls). There were no significant differences between the years of experience with the search systems (see Table 3). White and Morris (2007) have reported a correlation between the use of advanced search option and level of expertise. It was assumed that variations in technical search skills would not affect the search behaviour differently in the two groups, since there were no significant differences in formal training.

	Dyslexics		**Controls**	
Years of experience	***M***	***SD***	***M***	***SD***
General Web search engines	13.9	3.3	14.4	3.2
Bibliographic databases	4.6	3.9	3.9	3.1

Table 3: Years of search experience

The participants were asked about which Web search engines they usually searched. All 40 participants searched Google regularly. In addition, one dyslexic searched Safari, two dyslexics and one control used Bing and two controls searched Duck Duck Go. The students also listed the academic databases they used regularly.

A total of 11 dyslexics and 8 controls searched Bibsys Ask, while 4 dyslexics searched Google Scholar. In addition, 2 dyslexics searched the catalogues of the Norwegian Library of Talking Books and Braille. Furthermore, 11 dyslexics and 7 controls searched article databases through their libraries, such as ACM, PubMed and PsycINFO. Finally, one student in each group searched the public library catalogue Bibliofil. Consequently, all participants had previous experience with the search system applied in the Google experiment (EXP-II), but less experience with Bibsys (EXP-III).

Most students searched general Web search engines several times daily (see Table 4).

Search frequency	Dyslexics	Controls
Several times daily	85%	75%
Once a day	10%	20%
Once a week	0%	5%
Two-three times a week	5%	0%

Table 4: Search frequency Web search engines

A majority of the students searched bibliographic databases once a month or less (see Table 5). This was in contrast to their frequent use of Web search engines. This is in accordance with the findings by Jamali and Asadi (2010), who claim that Google Scholar has become the preferred information source for many scholars over bibliographic databases. Participants were expected to be more skilled in searching Google than Bibsys. However, there were no noteworthy between group differences, so this issue was not assumed to affect the results.

Search frequency	Dyslexics	Controls
Once a day	5%	5%
Two-three times a week	5%	0%
Once a week	20%	15%
Once a month	25%	5%
Rarely	30%	40%
Never	15%	35%

Table 5: Search frequency bibliographic databases

The students were also asked about the context of their searching (see Table 6). There was a clear distinction between the use of Web search engines and bibliographic databases, where the latter was mostly applied for educational purposes, while general Web search engines were used in several contexts.

	Web search engines		Bibliographic databases	
Context of use	Dyslexics	Controls	Dyslexics	Controls
Spare time	100%	100%	20%	20%
Educational purposes	95%	100%	95%	95%
Work	65%	60%	15%	5%

Table 6: Context of use

The participants were asked to self-rate their search skills in both Web search engines and bibliographic databases according to a four-levelled scale (see Table 7). No students rated themselves below average in Web search engines, although a majority of the dyslexics started by expressing that they were not really good at searching. However, a higher portion of dyslexics rated themselves as average compared to controls. This may be related to poor self-esteem which, according to Caroll & Iles (2006), is commonly found among many dyslexics.

The self-rating was more similar between participants concerning bibliographic databases, where a majority of the students were less confident.

	Web search engines		Bibliographic databases	
Self-rating	**Dyslexics**	**Controls**	**Dyslexics**	**Controls**
Excellent	40%	45%	5%	0%
Above average	25%	40%	15%	20%
Average	35%	15%	70%	65%
Below average	0%	0%	10%	15%

Table 7: Self-rating of search skills

Participants were asked about what (if anything) they found demanding when searching (see Table 8). Dyslexics regarded spelling search terms correctly as the most prominent obstacle. However, 8 of the 11 dyslexic students and one control who mentioned this spelling difficulty in EXP-II added that misspelling was not a problem due to the high tolerance for spelling errors in Google. However, these users often struggled with query formulations in other search systems.

All except one dyslexic emphasised the difficulties of spelling terms correctly in the bibliographic databases. In contrast, misspellings were not mentioned by any control users, who focused on the dissimilar ways of searching bibliographic databases compared to full text databases. Identifying proper search terms was regarded as difficult by both user groups. Regarding Web search engines, dyslexic users seemed most engaged in issues related to query formulations, while controls focused on assessment of result lists.

	Web search engines		Bibliographic databases	
Difficulties during search	**Dyslexics**	**Controls**	**Dyslexics**	**Controls**
Query formulation:				
Correct spelling of query terms	10	1	19	0
Identify proper search terms	3	3	2	2
Combine several terms into one query	1	0	0	1
Searching in English	2	1	2	0
Searching for very specific information	1	1	0	3
Searching peripheral topics	0	1	0	0
Searching for images	0	1	0	0
Search more "formally"	n/a	n/a	0	5
Match controlled vocabulary	n/a	n/a	1	1
Result lists:				
Select documents / find what you are looking for	1	4	0	3
English results lists despite Norwegian query	1	0	0	0
Irrelevant documents in the result lists	1	0	0	0
Keeping track of the documents	1	0	0	0
Retrieve updated documents	0	1	0	0
Too extensive result lists	0	0	1	1
Prefers Google Scholar over bibliographic databases	1	0	1	0

Table 8: Difficulties experienced during search (N=20 for both groups)

3.3 Testing participants

Five tests were conducted, with the purpose of obtaining indicators of visual and cognitive characteristics. The results were primarily used for screening purposes and comprised the following tests: a visual examination using Landolt C, a Word Chain Test to screen for dyslexia, a Corsi Block-Tapping Test and a Digit Span Test to assess short-term memory and a Stroop Test to map attention.

Both short term memory tests were conducted on a computer, using The Psychological Experiment Building Language (PEBL) and PEBL Test Battery (Mueller, 2010). PEBL is open-source software which includes several behavioural and psychological tests (Mueller & Piper, 2014). PEBL allows experimenters to alter the code, and thus enabled the translation of tests from English to Norwegian. This was necessary, since several dyslexics are reported to struggle with English language (Helland & Kaasa, 2005). The tests were administered in the same order for all participants at the first session in the lab, with standardised procedures and instructions (see Figure 9).

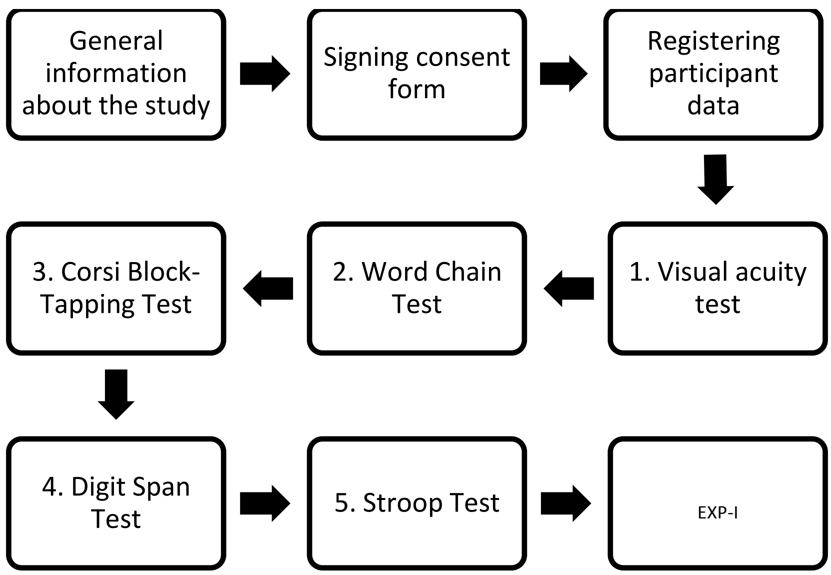


Figure 9: Sequence of session 1

All tests followed the acquired ethical guidelines and were approved by the Norwegian Social Science Data Service (NSD). The neuropsychological tests were administered according to the procedures described in Strauss, Sherman, and Spreen (2006), both regarding procedures and analyses. The tests were also discussed with a representative from the Norwegian Psychological Association (NPF) to ensure compliance with approved ethical procedures.

3.3.1 Visual acuity

Reduced visual function was expected to potentially affect the outcome of the experiments. Consequently, it was necessary to ensure that no participants had blurred or reduced vision. According to Ricci, Cedrone and Cerulli (1998), visual acuity testing is the most common assessment of visual function, and was therefore conducted for all participants.

Rabbetts (2007) describes visual acuity as a measurement based on the size of the smallest letter or symbol a participant can read from a test chart. Standardised symbols for testing vision, also referred to as optotypes, are applied in visual acuity tests, for instance Snellen and Landolt C (also called Broken ring). The Snellen chart includes letters; Landolt C involves a circle interrupted by a gap, one-fifth of its' outer diameter (ISO, 2009) (see Figure 10). There are no serifs on the Landolt C, and the borders of the brake are parallel.

Figure 10: The Landolt C optotype with different sizes and orientations. Image is retrieved from Wikimedia, and holds a public domain license

Colenbrader (2008) refers to Landolt C as a widely accepted reference optotype. According to Rabbetts (2007), Landolt C does not demand literacy and may therefore counteract decoding errors based on erroneous letter recognition. This view is supported by Reich and Ekabutr (2012), who reported that Landolt C is often used to test illiterate persons.

The Landolt C optotype may be displayed in different sizes depending on the distance of the testing, and usually has 4 or 8 orientations (ISO, 2009). In this study, charts with four orientations were used: gap up, down, right or left, and participants were instructed to indicate the direction of the gap. Although some dyslexics are reported to confuse left and right (Stein, 2001), this was solved by indicating the orientation of the C by pointing or relating to physical features in the lab.

Two standard Landolt C visual acuity charts from Precision Vision were used at standard measuring distances of 4 meters for distance and 40 cm for near vision (ISO, 2009). The optotypes were displayed in black ink on a white background. The distance vision chart measured 23 x 35.5 cm. The chart had two smaller groups for repetitive testing (see Figure 11). The participants were asked to follow the left path; the right was used if participants seemed to remember the combinations from previous repetitions. The chart was wall-mounted and a marker was placed on the floor to ensure an accurate distance between the participant and the chart.

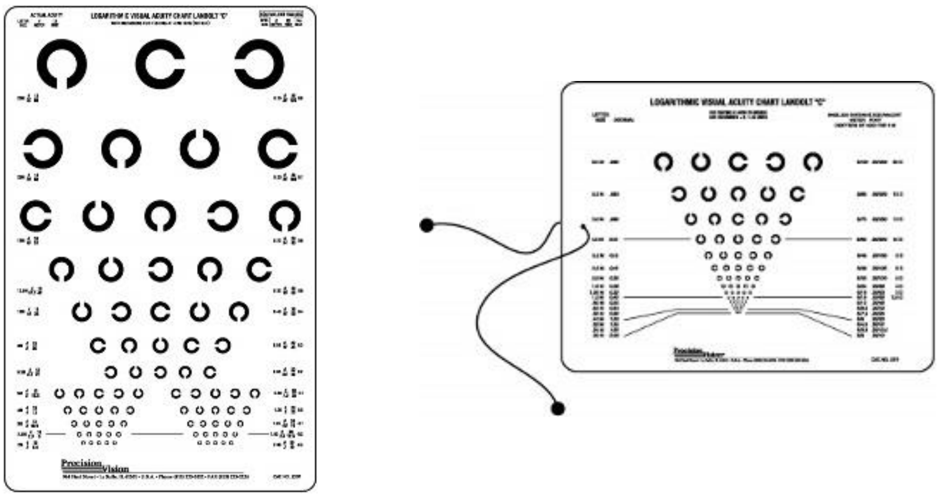


Figure 11: Distance chart (left) and near chart (right), image courtesy of Precision Vision

The near vision chart measured 18 x 23 cm. A string fastened to the chart ensured an accurate viewing distance (see Figure 11). The participants were sitting with the chart in their hands, in the same position as they would place a book when reading.

The visual acuity testing followed the standardised procedure for all participants described by Ferris and Bailey (1996): First distance vision testing (right eye, left eye, binocularly), then near vision (right eye, left eye, binocularly). A vision occluder was used as cover when examining one eye, to block vision and avoid light from reaching the other eye. A total of 14 participants wore glasses or contact lenses (nine dyslexics and five controls). Because these visual aids were used during the experiments, glasses or lenses were kept on during the visual examinations.

All the included participants had normal or normal to corrected vision, with at least an acuity of 0.6 on each eye separately and 0.8 with both eyes open, at both tests. Two dyslexic users were excluded from the experiments due to low scores on the visual acuity test for near sight, with values from 0.32-0.50 on one or both eyes respectively.

3.3.2 Dyslexia screening test

A screening test for dyslexia was conducted for both participant groups. Although the dyslexic students reported to have been previously diagnosed by psychological or pedagogical professionals, the test eliminated the need for consulting sensitive medical records to confirm the diagnosis. According to Mortimore and Crozier (2006), students may not be aware of their dyslexia. Consequently, the comparison group was tested to ensure than no controls were dyslexic.

The word chain test is based on word recognition. Such a test may not be used for diagnostic purposes, but provides indications of dyslexia, and is therefore suitable for screening purposes. Word chain tests are quick and easy to administer and has been used in previous research in several versions and languages, for instance Wolff and Lundberg (2002), Samuelsson, Herkner and Lundberg (2003), Wolff and Lundberg (2003), Solheim and Uppstad (2011) and Dahle and Knivsberg (2014).

The word chain test applied in this study was Ordkjedetesten (translates The Word Chain Test), designed for Norwegian language by Høien and Tønnesen (2008). It is based on a test developed in 1987 as a part of a multidisciplinary research project on reading development in Sweden to map decoding skills and word recognition. In this study, the 5th version in bokmål (one of two standard forms of the Norwegian language) was used, produced by Logometrica (Høien & Tønnesen, 2008), from now one referred to as The Word Chain Test. Word chain tests exist for other languages as well, for instance Swedish and English (Guron, 1999).

The Word Chain Test (Høien & Tønnesen, 2008) consists of 90 word chains. Each word chain contains four relatively common words in a sequence, where the inter-word spaces are omitted. Word lengths vary from two to seven letters, and the chains consist of nouns, verbs, adjectives, adverbs, prepositions and names of digits. The participant is asked to mark the lacking inter-word spaces, three for each chain. Each word chain is awarded one point, given that all the three boundaries are marked correctly (see Figures 12 and 13). If errors occur, zero points are allocated that specific word chain. The test has a maximum score of 90 points.

The test comprised three pages of white paper with 30 word chains printed in black ink on each page. The participants were asked to indicate the boundaries between the words with a pen (see Figures 12 and 13) in as many word chains as possible within a four-minute time limit. Ahead of the test, the participants had received instructions, and completed a practice sheet containing six word chains. The test was timed using a stop-watch and the experimenter sat in another part of the room to give the participants peace and quiet during the test.

Figure 12: The Word Chain Test, line 8, dyslexic user, scored 1/3 points
Reproduction permitted by Logometrica

Figure 13: The Word Chain Test, line 8, control user, scored 3/3 points
Reproduction permitted by Logometrica

The Word Chain Test contains normative scores for adults, based on 421 participants, with a mean value (*M*) of 55.2 points and a standard deviation (*SD*) of 11.8 points (Høien & Tønnesen, 2008). A low score is considered indicative of dyslexia, and diagnostic tests are recommended for adult scores below 43 (Høien and Tønnesen, 2008). As a part of the preparations for the study, the experimenter piloted The Word Chain Test on users with and without dyslexia who were not participating in the study. The results were in accordance with the standardised scores and the piloting provided the experimenter administering practice.

The results from the Word Chain Test confirmed group affiliation. The users with dyslexia obtained significantly lower scores (M=39.7, SD=10.2) compared to the control group (M=60.3, SD=9.7), $t(40)$=6.7, p<.001, d=2.1.

3.3.3 Visuospatial short-term memory

The Corsi Block-Tapping Test assesses visuospatial short-term memory (Vandierendonck, Kemps, Fastame & Szmalec, 2004). According to Pagulayan, Busch, Medina, Bartok and Krikorian (2006), this test is widely used. The original test displayed nine identical blocks irregularly positioned on a wooden board. The experimenter pointed to one block per second in a specific order, and the participant was asked to reproduce the sequence. The sequence lengths increased until the recall was incorrect (Vandierendonck et al., 2004). Corsi Block-Tapping Tests may be conducted forwards or backwards, and with crossing sequences or non-crossing sequences. According to Vandierendonck et al. (2004), performance is better in the forward recall tasks than the backwards.

In contrast to the original Corsi Block-Tapping Test, The PEBL version applied in this study was conducted on a computer. According to Vandierendonck et al. (2004), computerised versions of the Corsi Block-Tapping Test are now commonly applied. The computerised version ensured identical experimental procedures and instructions for all participants.

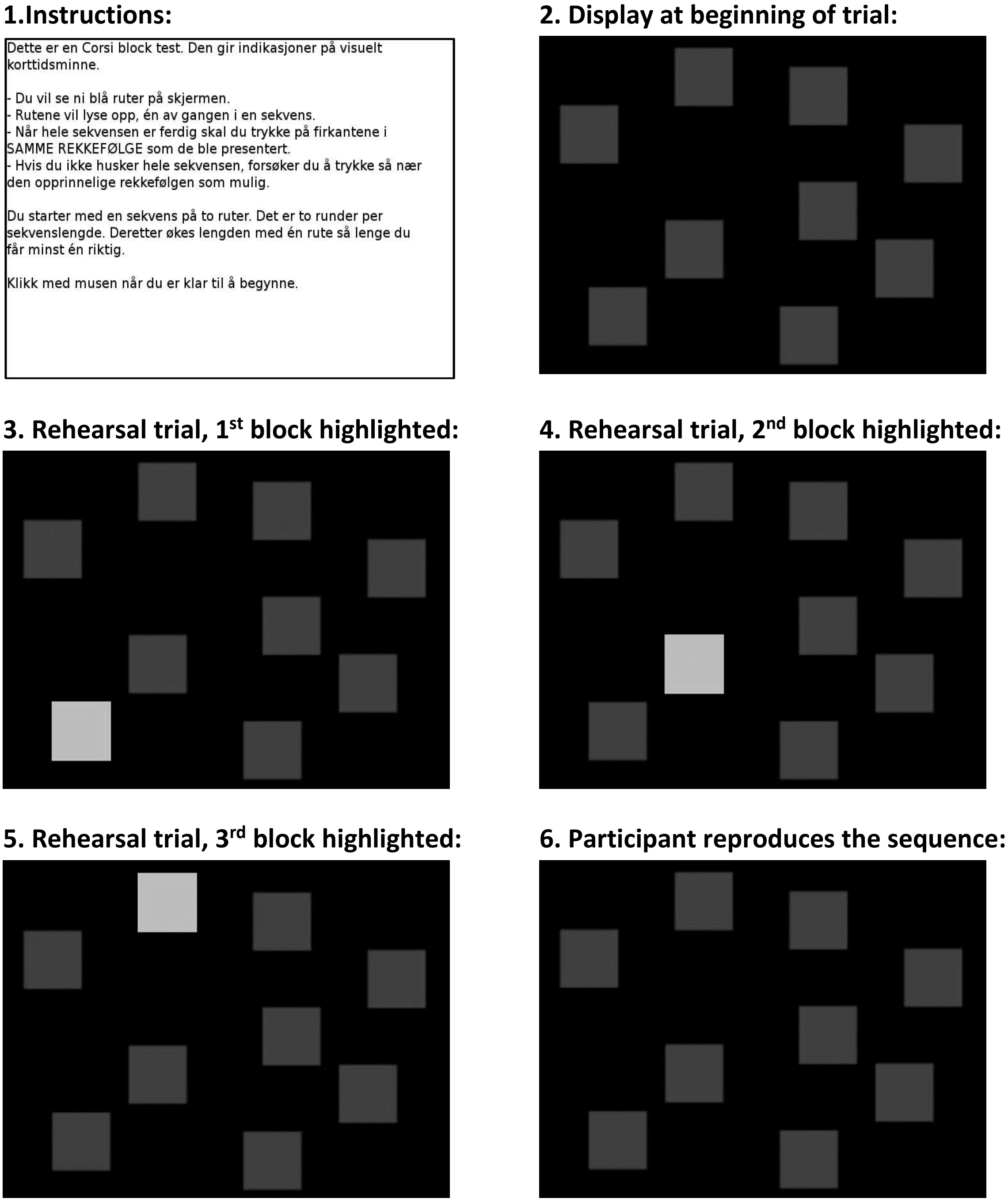


Figure 14: Example of Corsi Block-Tapping Test, rehearsal trial with a sequence length of three blocks

In the PEBL version, the user is first presented with instructions, and three rehearsal trials. For each task, a black screen appears; displaying nine navy blue blocks irregularly laid out on the screen in fixed positions (see Figure 14). For each trial one block at a time is highlighted with yellow colour in a sequence, and the task is to reproduce the sequence of the blocks when the whole sequence is finished. The test starts with a sequence of two highlighted blocks. The sequence lengths increase with one block for every second trial, and the test is completed when the participant fails to recall both sequences with the same length.

The Corsi Block-Tapping Test is reported to be widely applied in clinical research (Pagulayan et al., 2006; Kessels, Zandvoort, Postma, Kappelle & Haan, 2010). A variety of versions and administration procedures have been used, and Pagulayan et al. (2006) report that it has been difficult to develop normative data. Kessels et al. (2000) suggest a mean block span of 6.2 (*SD*=1.3) and a mean score of 55.7 (*SD*=20.3) for healthy controls. Pagulayan et al. (2006) tested school children and young adults, and found a large effect for grade. The young adults had a mean block span of 7.1 (*SD*=1), and it was assumed that spatial span may not increase after age 14.

The Corsi Block-Tapping Test has been conducted on dyslexic users by several researchers. Some studies on dyslexic children concluded that were no significant differences between dyslexics and controls (Gould & Glencross, 1990; Palmer, 2000). Jeffries and Everatt (2004) found lower block spans among children with dyslexia compared to controls, but these differences were not significant. Smith-Spark and Fisk (2007) assessed 22 dyslexic students and 22 controls, and reported that the dyslexics had a lower span than the controls, but the differences fell just short of being significant.

Dyslexic participants achieved significantly lower block spans (M=5.52 blocks, SD=0.93) than controls (M=6.43 blocks, SD=1.33), $t(40)$=2.56, p<.015, d=0.81, which indicates a reduced visuo-spatial short-memory capacity. These results are in contrast to previous findings, that people with reading disorders have an unimpaired performance on Corsi tasks. However, the findings are in accordance with impairments in the short-term memory capacity documented in dyslexics (Smith-Spark & Fisk, 2007).

3.3.4 Verbal short-term memory

The Digit Span Test is a widely used method for assessing verbal short-term memory, which Helland and Asbjørnsen (2004) have described as the most commonly referred marker for dyslexia. A computerised, forward open-source software version in PEBL (Mueller, 2010) was used. The Digit Span Test may also be conducted backwards, but is then regarded as a working memory task (Helland & Asbjørnsen, 2004).

In the PEBL version, the participant is first presented with instructions (See English version in Figure 15). Instructions were translated to Norwegian for test purposes, since dyslexics are reported to struggle with English language (Helland & Kaasa, 2005). The instructions are followed by the sequence length of the first string. For each trial, the numbers are presented on screen as a digit, one at a time, and the participant has to recall and input the whole string in the correct order. The sequence length increases with one digit for every second trial, and the test is completed when the participant fails to recall both strings with similar length.

Dyslexic students had a significantly lower memory span (*M*=5.29 digits, *SD*=1.06) compared to the control group (*M*=6.14 digits, *SD*=1.20), $t(40)=2.46$, $p<.02$, $d=0.78$. This is in accordance with for instance Gould and Glencross, (1990), Helland and Asbjørnsen (2004), Jeffries and Everatt (2004) and Smith-Spark and Fisk (2007).

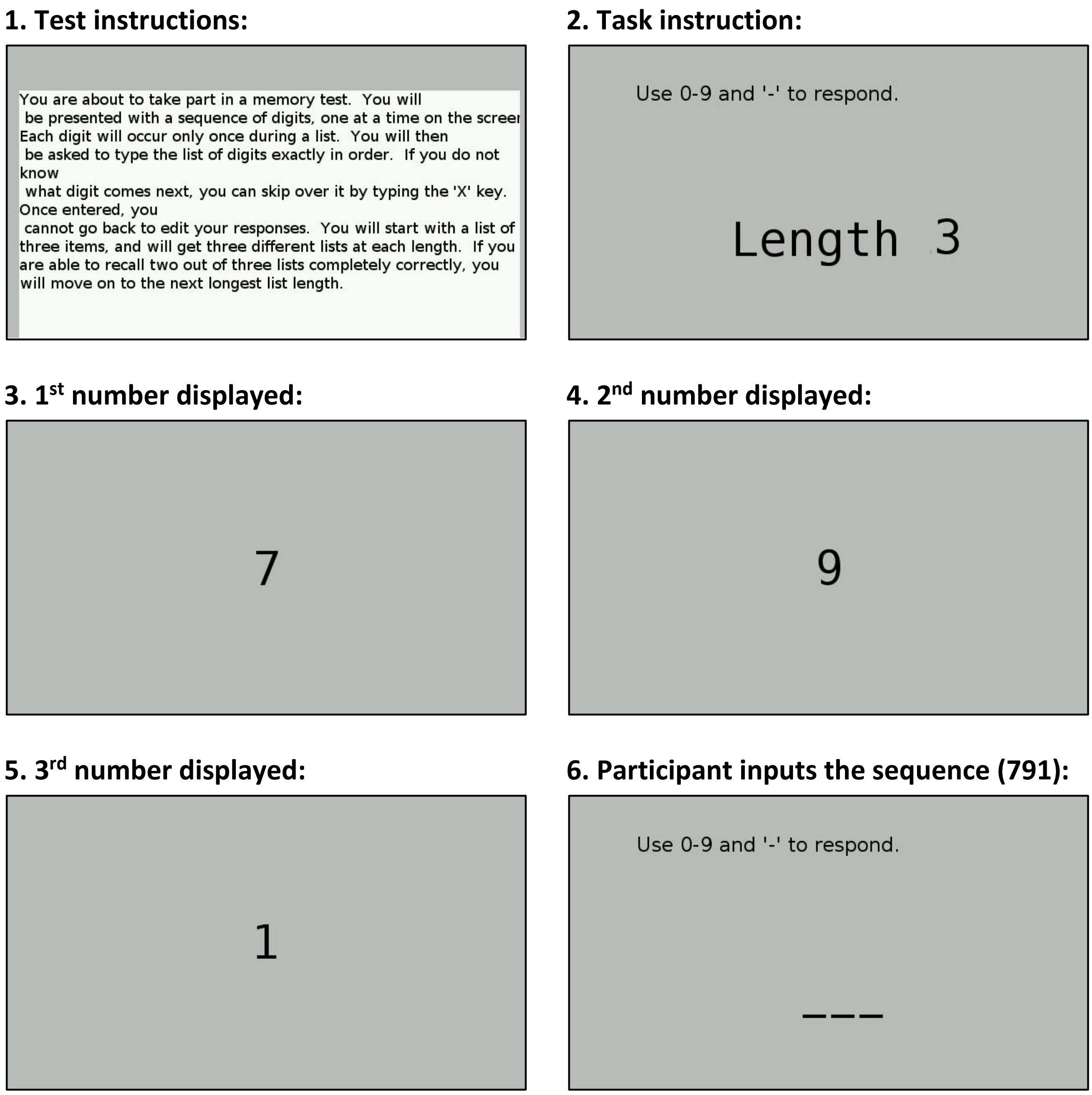


Figure 15: An example of the Digit Span Test, reproduced from http://pebl.sourceforge.net/screenshots.html. All the images are released into the public domain. A Norwegian translation was used in the study.

3.3.5 Attention

The Stroop Test, also referred to as the colour-word-interference test, is widely used for assessing selective attention. The purpose is to reveal for how long a person can maintain a goal in mind and at the same time suppress a habitual response (Strauss et al., 2006). The Victoria Stroop Test was applied, among others because it is not as extensive as the original test, which makes the administration time considerably shorter and the respondents less exhausted. Further, Strauss et al. (2006) found that if a person struggles with the tasks, this will become most evident if the test is short, among others due to lack of practice. Another advantage with the Victoria Stroop test is that it is released in the public domain.

The Victoria Stroop Test consists of three test cards (see Figure 16-18); Dot-card, Word-card and Colour-card. Test cards were created according to the instructions in Strauss et al. (2006). The dot-card consists of 24 dots in four different colours (green, blue, yellow and red). Each colour is presented once per row. The Word-card displays 24 common words, printed in four different colours. The Colour-card contains the names of the four colours. However, the word is never printed in ink which corresponds to the colour the word refers to. For instance, the word Blue is printed in green, yellow or red ink, but never blue.

A Norwegian official version for the Word-card was not available, so four proper Norwegian words had to be selected. The English version included the words *and, when*, *over* and *hard* (Strauss et al., 2006). The Victoria Stroop Test also exists in a French version using the words *pour* (for), *mais* (but), *quand* (when) and *donc* (so) (Bayard, Erkes & Moroni, 2009). Both the English and French version include three words from the hundred most frequently used words in their language (Oxford Dictionaries, n.d.; LanguageDaily, n.d., see Table 9).

English	Frequency	French	Frequency
and	5	pour	29
when	51	mais	35
over	78	quand	74
Hard	-	donc	172

Table 9: Ranking on word frequency lists

In the English version, *when* and *and* represent conjunctions, *over* is a preposition and *hard* is an adjective or adverb. The French version includes three conjunctions and one preposition. Three conjunctions and one preposition extracted from a Norwegian word frequency list were used (Korrekturavdelingen, n.d., see Table 10): *og*, *for*, *men* and *når* (translates and, for, but, when). Three of these words are included in the French version; two in the English version (see Table 10).

Norwegian	Frequency	English version	French version
og	1	✓	×
for	12	×	✓
men	23	×	✓
når	53	✓	✓

Table 10: Norwegian words used in the Word card

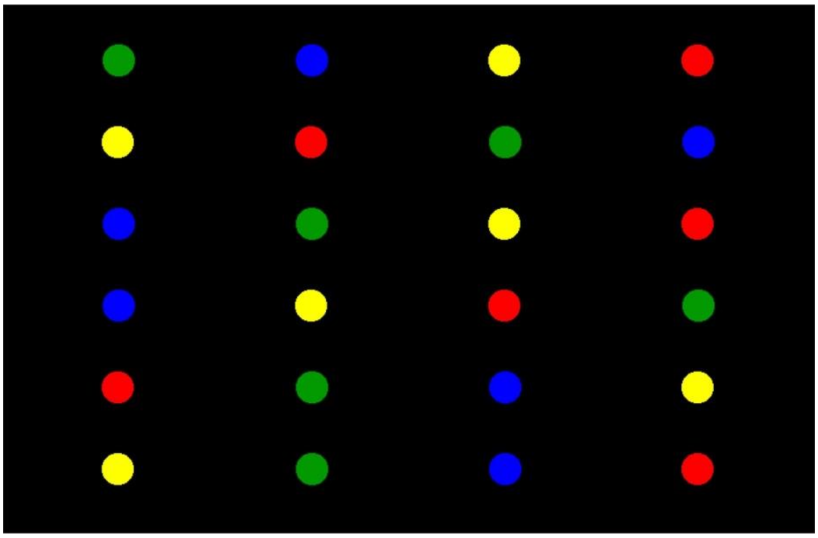


Figure 16: Dot-card

Figure 17: Word-card

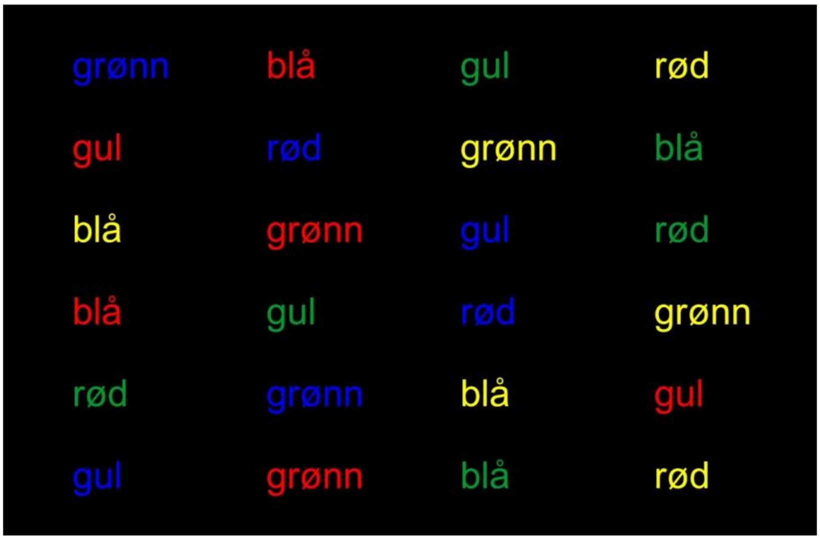


Figure 18: Colour-card

The testing followed standard procedures, as described in Strauss et al. (2006). First, participants were given the Dot-card (see Figure 16) and instructed to read out the colours of the dots. On the Word-card task (see Figure 17), the participants were asked to read the colour of the ink, and disregard the verbal content. On the Colour-card task (see Figure 18), respondents were asked to read the colour of the ink and again ignore the verbal content.

Participants were instructed to read as fast and correct as possible, and were corrected if they made errors. Spontaneous corrections were regarded as correct, according to standard procedure (Strauss et al., 2006). An audio recorder was applied during the test, and a stop watch was used during analysis. Time usage and number of errors were recorded for each card.

A commonly used measurement of the Victoria Stroop Test is a difference score between the time required to read the Colour-card versus the Dot-card. High scores are considered indicative of a large interference (Strauss et al., 2006). Although widely used, Bayard, Erkes and Moroni (2011) have reported that few normative data for the Victoria Stroop Test exist. Moreover, according to Malek, Hekmati, Amiri, Pirzadeh and Gholizadeh (2013) using norms of other languages may be erroneous, due to differences in accents and the speed of articulation. The main purpose in this context was to compare the two participant groups. Consequently, the difference between the groups was most significant.

The students with dyslexia had significantly higher difference scores on the Stroop-test (M=15.2 points, SD=15.1) than controls (M=8.10 points, SD=3.45), $t(40)=2.10$, $p<.043$, $d=0.7$. This was expected, since larger Stroop effects are documented among dyslexics than non-dyslexics (Faccioli, Peru, Rubini & Tassinari, 2008). According to Helland and Asbjørnsen (2004) the Dot-card may be used to assess naming speed. Dyslexic students took significantly longer on the Dot-card task (M=16.4 s., SD= 6.2) compared to controls (M=12.7 s., SD=2.2), t(40)=2.62 , $p<.013$, $d=0.8$. This is in accordance with previous findings, that dyslexics are impaired in rapid naming tasks (Lervåg & Hulme, 2009).

3.4 Experimental design

All the three experiments were conducted during one academic year. The first experiment (EXP-I) started in July 2013, the second experiment (EXP-II) in January 2014 and the third experiment (EXP-III) in March 2014. Ideally, the experiments could have been distributed over a longer time span, which would have enabled analysis of one experiment before designing and conducting the next. However, such a design would have increased the risk that participants would withdraw from the study, finish or quit studying or move and thus not be available, or no longer comply with the inclusion criteria of the study. Furthermore, it was an advantage that the participants did not have any significant changes in search experiences or training between EXP-II and EXP-III.

The first experiment was carried out in the same session as the screening tests and cognitive tests described in Section 3.3. However, the schedule was laid out so that the participants would not be exhausted when starting EXP-I, and included a break before initiating EXP-I. The experimental setup was identical for all the experiments, as described in Section 3.1.2. According to Gross' (2005) theory on The Imposed Query (see Section 2.4.3), the quality of question transfer and characteristics of the agent/imposer affect the outcome of imposed queries. Consequently, the instructions and presentation of tasks were given much attention and piloted before the experiments were carried out.

3.4.1 EXP-I – Result lists

The purpose of the first experiment (EXP-I) was to study search performance and search preferences in layouts containing either icons, words or both in dual-modality displays, to investigate whether graphic content in result lists may improve result list assessment among dyslexics (H_1). Four different task types were presented (see Figures 19 – 22): icon arrays (ICON), word arrays (WORD), icons and words in a grid-layout (ICON/WORD-GRID) and icons and words in a list-layout (ICON/WORD-LIST).

The one-modality tasks had a set-size of 24 objects; one target and 23 distractors laid out in a grid of 6x4 (6 columns, 4 rows). The dual-modality displays contained 48 items; containing 24 icons and 24 words representing the same noun laid out in a grid of 6x8 or a list of 2x24. Consequently, there were two target items and 46 distractor objects in the dual-modality displays, and the task was solved by locating either the word or the icon.

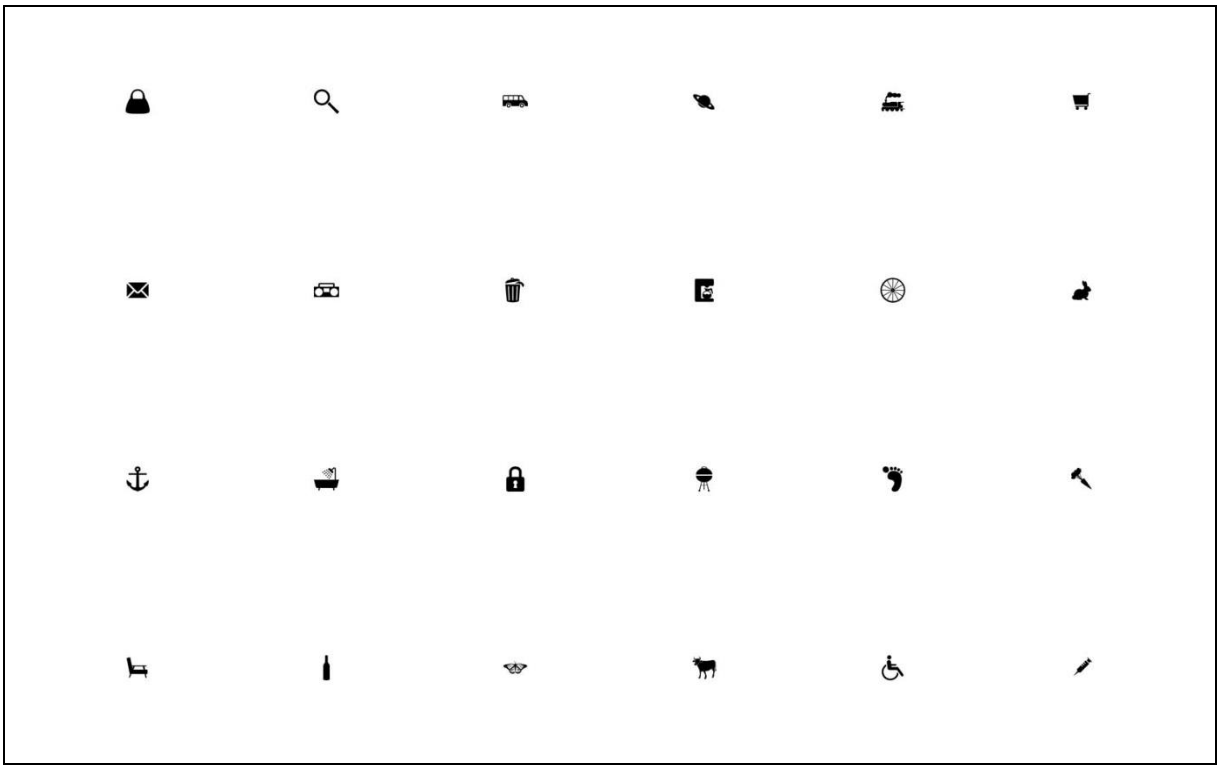


Figure 19: Example stimulus ICON task EXP-I

Kopp	Is	Sol	Bjelle	Glass	Stol
Symaskin	Timeglass	Klokke	Koffert	Kvinne	Gulrot
Buss	Sommerfugl	Veske	Jordbær	Barnevogn	Hamburger
Søppelbøtte	Bamse	Kryss	Nøkkel	Sky	Grill

Figure 20: Example stimulus WORD task EXP-I

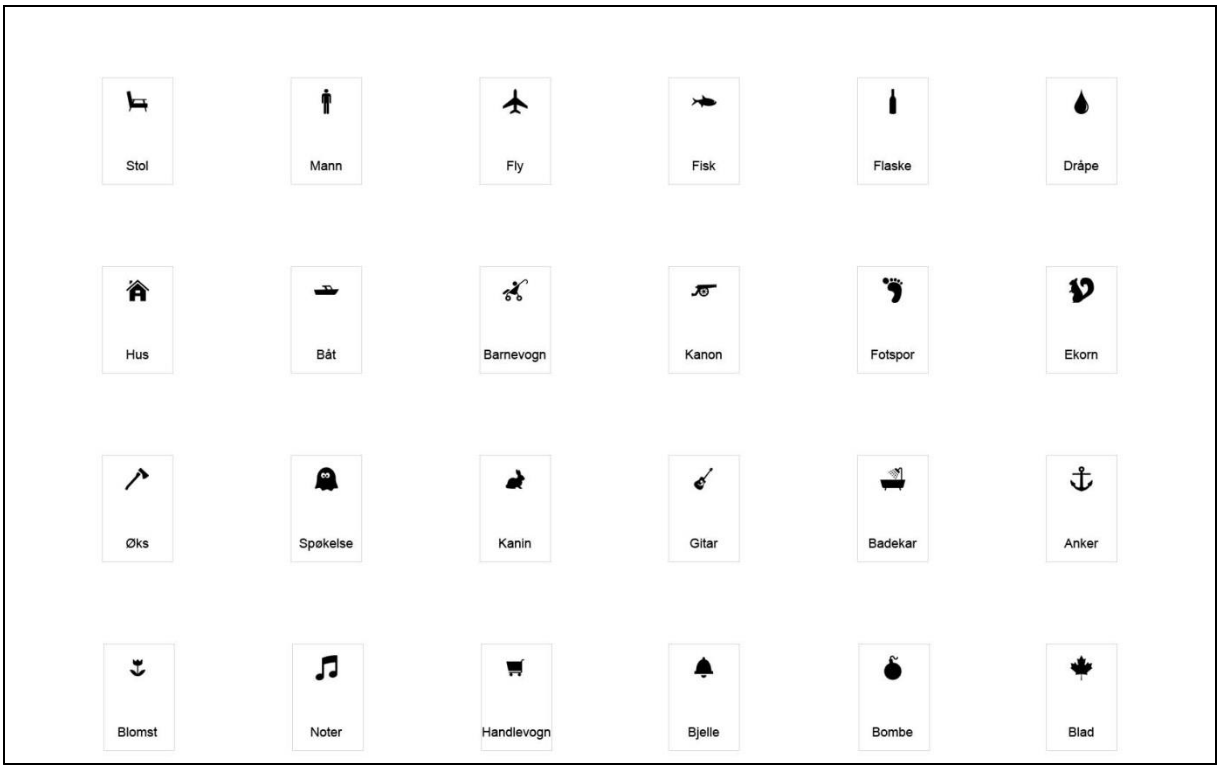


Figure 21: Example stimulus ICON/WORD-GRID task EXP-I

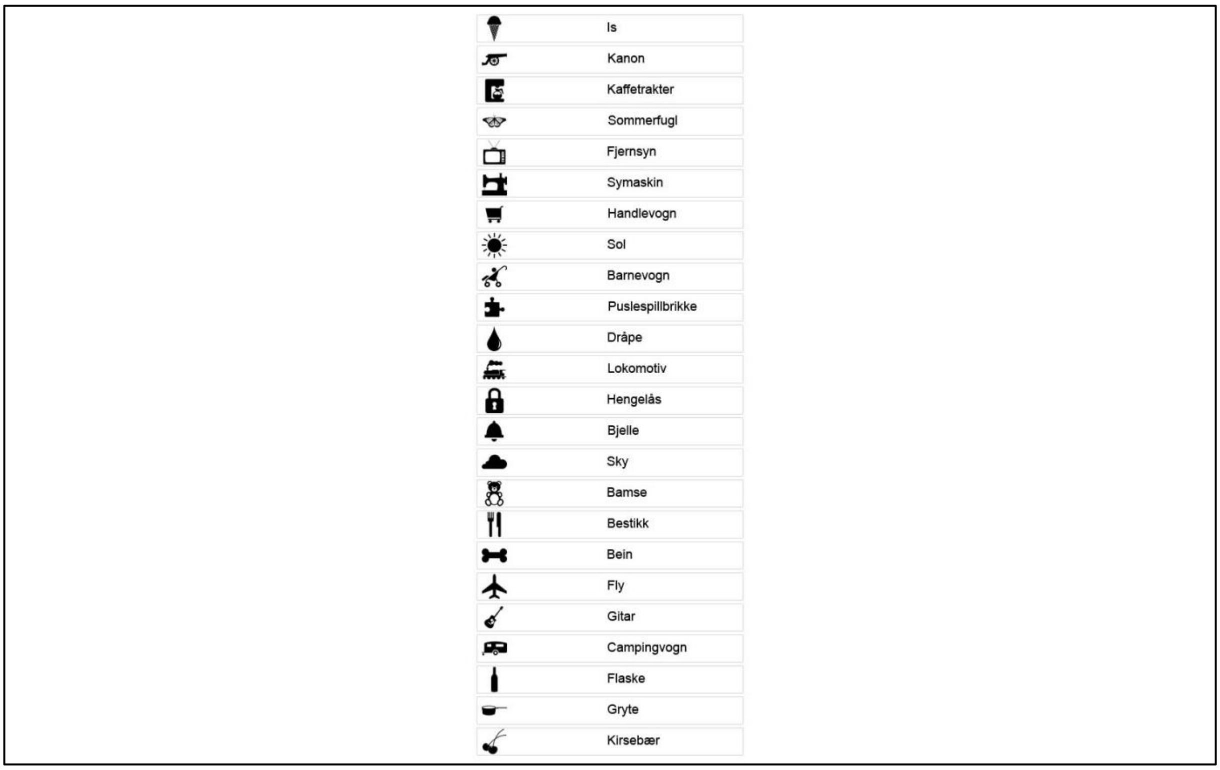


Figure 22: Example stimulus ICON/WORD-LIST task EXP-I

All tasks followed the same procedure displaying three screens (see Figure 23). First, the participants were presented with a blank screen and pre-recorded verbal instructions on which object to search for. Instructions were followed by a fixation cross, which was designed as an AOI (Area of Interest) trigger, entailing that participants had to look directly at the fixation cross for 1 second before the actual task was displayed. This ensured that all participants fixated on the same location on the screen when the search task was initiated, and allowed for scanpath comparisons. Such a design has also been applied in other studies, for instance Peterson et al. (2001) and Becker (2011). When the participants had located the target, they were instructed to fixate on the object and push the space bar.

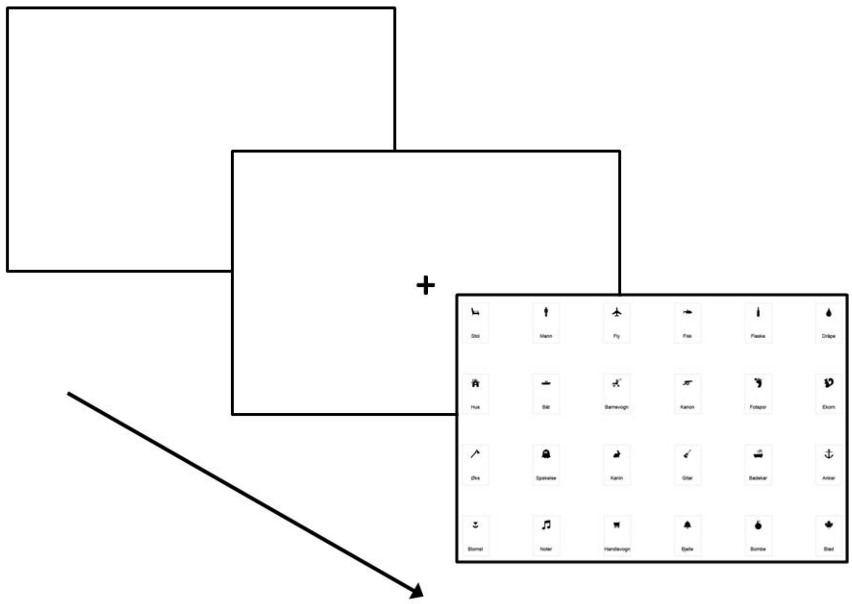


Figure 23: Procedure EXP-I (Berget, Mulvey & Sandnes, 2016)

Icons were downloaded from the Noun Project (http://thenounproject.com) and were released into the public domain under a Creative Commons license (see Figure 24 for examples of icons). All the icons represented nouns, and were screened by 64 students during pilot testing who were not participating in the actual experiments. However, four questionnaires were discarded because several fields were not completed. The students were given a paper based questionnaire including 105 icons, and asked to write down a word to describe what the icon represented. To prevent errors caused by misinterpretations, only icons with a rater agreement of 100% were used as targets in the search tasks.

Differences in interpretations related to gender and age among the participants are discussed in Berget and Sandnes (2015b), see Appendix E. Icons with a rater agreement below 90% were removed entirely from the icon bank. The remaining icons were used as distractors. After the exclusion of low rated icons, the icon bank consisted of 99 icons. The icons were presented in black on a white background, in a size of 50 x 50 pixels (equalled approximately 1° of vision on screen) with a resolution of 72 pixels/inch.

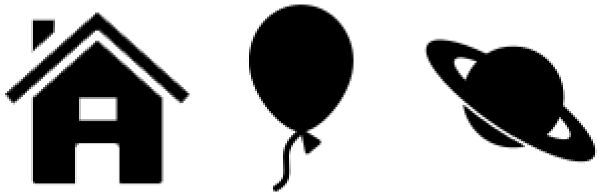


Figure 24: Examples icons EXP-I

The textual stimuli contained common Norwegian words equivalent to the nouns represented by the icons, such as House, Balloon and Planet for the icons presented in Figure 24. The word lengths were between 2 and 21 letters in Arial black font, 18 points, displayed on a white background. It was assumed that short and long words would stand out from the remaining words. Consequently, only words with intermediary visual word lengths were included as targets to avoid that word lengths would create a pop-out effect, and thus affect search performance, as described by Nothdurft (1999).

All objects were randomly extracted from the icon/word-bank. An item could appear in multiple tasks, but did not reappear after functioning as a target. This was done to avoid that dyslexic participants should mix up tasks due to a reduced short-term memory capacity. However, this experimental design led to a fixed order for tasks for all participants, which may be regarded as a limitation of the study. However, it was regarded as a potentially bigger problem that tasks would have to be discarded due to participants locating the wrong targets or taking much time looking for the wrong target.

The size of the icons/words and the object spacing was piloted multiple times to ensure that only one item could be processed per fixation (see Section 2.5.2 on the relationship between eye movements and attention). The list-layout had a decreased object spacing, to avoid reducing the object sizes or set-sizes.

EXP-I was initiated by one rehearsal task for each task type, followed by sections of six tasks of each type. Every task appeared in two sections in the same order for all participants (see Figure 25).

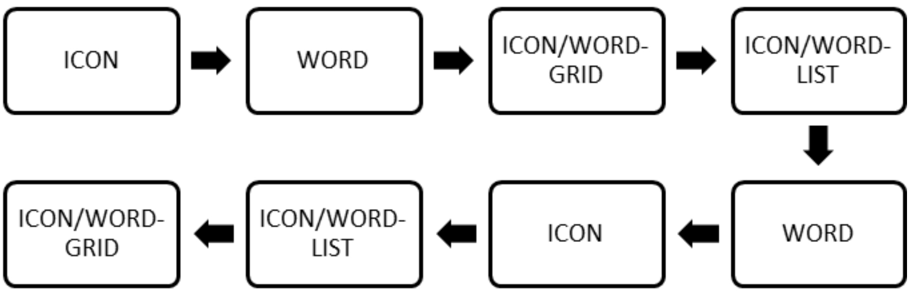


Figure 25: Task order EXP-I

3.4.2 EXP-II & EXP-III – Query formulation

The experimental designs for EXP-II and EXP-III were identical, except for different search systems and tasks. This section describes the common experimental design features and procedures, while details concerning each experiment are presented in Sections 3.4.3 and 3.4.4.

Two sets of ten predefined search tasks were applied (see Table 11 for EXP-II and Table 12 for EXP-III). The tasks were solved in the basic search user interfaces of Google for EXP-II and Bibsys Ask for EXP-III. The participants were not provided with any search instructions, except for formulating queries in Norwegian, to enable query comparisons in the subsequent data analysis. Further, English was avoided because many Norwegian dyslexic users consider English difficult, according to Helland and Kaasa (2005). No training or rehearsal trials were given, since all the students had prior experiences with the search user interfaces.

The tasks were presented verbally, to avoid misinterpretations due to word decoding errors and to prevent participants from seeing the spelling of the words. An image representative of the question was displayed during instructions (see Figure 26) and the task number allowed the participants to track the progress of the session.

Figure 26: Example stimulus task 10 EXP-III

The tasks were presented in a fixed order for all participants. Experimental task orders are commonly randomized to counterbalance learning effects. However, in this study the most important design consideration was not to make the dyslexic students feel uncomfortable or give up during the session. It was assumed that the combination of several subsequent difficult tasks and poor self-esteem issues (Caroll & Iles, 2006) could result in dyslexic participants quitting before completing the experiment.

Several dyslexics also expressed discomfort with testing situations. Consequently, the tasks were presented in a fixed order of increasing difficulty. Moreover, the first and middle tasks were easier to solve, to encourage the participants at the beginning and to provide a feeling of coping during the experiment. All the students had prior experience with the search systems (see Section 3.2.3), and it was thus assumed that learning-effects would be minimal. Further, the search sessions usually took less than 15 minutes, so no major fatigue-related effects were expected.

The tasks had a moderate complexity level, and could be solved by one relevant document. Participants were asked to not terminate the search process until an answer was found. All the tasks concerned topics not directly related to the field of study for any of the participants, to ensure that differences in background knowledge would not affect the outcome.

The search tasks were not based on a collection of documents with a set of predefined queries, which is often used as an evaluation approach to evaluate search systems (Borlund, 2003). The tasks were constructed for this particular study. Borlund (2003) claims that tasks should be tailored according to the information environment and the test persons. For instance, the situations should be possible for the participants to relate to, should be topically interesting and provide enough context for the users to relate to (Borlund, 2003).

The tasks in this experiment were constructed to provoke spelling errors which are typically made by dyslexics. The tasks included compound words: kildesortering, double consonants: Oppland, consonant clusters: bakovervendt, silent consonants: Sigrid (d is not pronounced), silent vowels: Algerie (e is not pronounced) and words with irregular orthography, where the spelling and pronunciation differs: Shakespeare. Such words are typically reported to be frequently misspelled by dyslexic users (Helland & Kaasa, 2005; Lyster, 2007; Høien & Lundberg, 2012). Consequently, the query terms were regarded more difficult to spell for dyslexic users than controls.

The sessions followed the same procedure for all participants, starting with pre-interviews and general instructions, followed by the information retrieval experiment and ending with a debriefing. The results from the pre-interviews are presented in Section 3.2.3. During the search sessions the participants were not spoken to unless they asked a question to the experimenter.

3.4.3 EXP-II – Web search engine

The second experiment was designed with a twofold purpose; to study the use of query-building aids (addressed by H_2) and to observe how spelling difficulties affected the search performance in a search system with a high tolerance for spelling errors (related to H_3).

EXP-II was conducted in the general Web search engine Google, which is widely used (McCallum & Bury, 2013) and has a high tolerance for spelling errors. If a user misspells a query, Google typically displays a result list for what is assumed the correct query (see Figure 27). Further, there are built-in query-building aids, such as an autocomplete function (see Figure 28) and a red line which is displayed in the search box beneath assumedly misspelled words.

Figure 27: Spelling correction in Google. Google and the Google logo are registered trademarks of Google Inc., used with permission.

Figure 28: The autocomplete function in Google. Google and the Google logo are registered trademarks of Google Inc., used with permission.

The basic search user interface in Google comprises one search box and a search button. Google is a full text database, and the query terms are therefore compared with the whole text of each document. Full text systems require queries formulated in natural language, but may also demand the application of synonyms to increase recall (see Section 2.4.1). However, since these tasks were not exhaustive tasks, where several documents are needed to fulfil the information need, a comprehensive result lists with a high recall were not required for a successful outcome. Consequently, the participants did not have to include synonyms in their queries.

The predefined tasks applied in EXP-II are presented in Table 11, alongside suggested query terms. The expected query terms in Table 11 should merely be regarded as suggestions of concepts which may be included in the queries to solve the tasks. The queries could be more precisely or more widely defined than suggested below. For instance, task three could have included the term album, but was not necessary to solve the task. Participants were given unlimited search time.

Task	Task (originally given in Norwegian)	Expected query terms	Translated query terms
1	Find the homepage of the Norwegian royal family?	Kongefamilien	The Royal Family
2	Until which age should children sit in a rear-facing car seat?	bakovervendt bilsete alder	rear facing car seat age
3	Which album by Bob Dylan was released in 1969?	Bob Dylan 1969	Bob Dylan 1969
4	What is the size of the targets in prone position in biathlon?	skiskyting blinkstørrelse	biathlon target size
5	What are the rules for alcohol consumption and boating?	alkohol båtkjøring	alcohol boating
6	Who has written the books about Curious George?	Nysgjerrige Nils	Nysgjerrige Nils
7	How should private households sort their garbage in Drammen?	kildesortering private hus Drammen	recycling rubbish private households Drammen
8	Who is the voice of Pelle Politibil in the animated movies?	Pelle Politibil stemme	Pelle Politibil voice
9	How many people live in Jamaica?	befolkningstall Jamaica	population Jamaica
10	In which museum in Egypt can you find the famous gold mask of Tutankhamun?	Tutankhhamon gullmaske museum Egypt	Tutankhamun gold mask museum Egypt

Table 11: Predefined search tasks EXP-II

3.4.4 EXP-III – Library catalogue

The purpose of EXP-III was to investigate the effect of dyslexia in a search system with a low tolerance for spelling errors (H_3), namely the Norwegian library catalogue Bibsys Ask. This search system has no query-building aids, and could therefore not contribute to answer H_2. The participants were allowed to use other online resources as supplementary aids, but were encouraged to first try and solve the task in Bibsys Ask. Further, the final answer to the task had to be located in Bibsys Ask. The searches in external resources applied query-building aids and are therefore included in the discussion of H_2.

The Norwegian version of the simple search user interface in Bibsys Ask was used (see Figure 29). Bibsys Ask is applied by higher education libraries in Norway, in addition to several research libraries and institutions. Bibsys Ask has a low tolerance for spelling errors (see Figure 30), and the only search assistance is an authority file. Authority files include the authorised form of certain words, for instance author names and keywords.

Bibsys Ask is a bibliographic database, and does not offer full text search. Queries are compared to manually indexed metadata, such as author names, book titles and subject terms. The result list does not contain the actual documents, but document surrogates, which merely describe the document.

Figure 29: Simple search in Bibsys Ask, reproduced with permission.

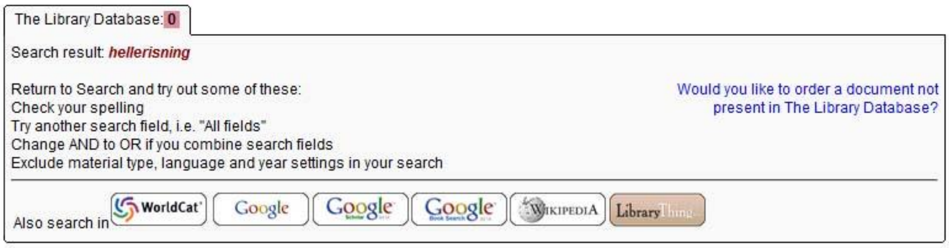


Figure 30: User interface in Bibsys Ask succeeding an erroneous query, reproduced with permission

The tasks for EXP-III are presented in Table 12, including suggested query terms. Query terms were more clearly included in the tasks in EXP-III than EXP-II due to higher demands for correct query terms in bibliographic databases compared to full text databases. Further, the tasks in EXP-III were more directly focused on specific authors and book titles, since books represent the predominant content type in Bibsys Ask.

Task	Task (originally given in Norwegian)	Expected query terms	Translated query terms
1	Find a document about Vigeland Sculpture Park?	Vigelandsparken	Vigeland Sculpture Park
2	In which year was the book "Rock carvings in Hedmark and Oppland" published?	helleristninger Hedmark Oppland	rock carvings Hedmark Oppland
3	Find a book written by Sigrid Undset about Kristin?	Sigrid Undset Kristin	Sigrid Undset Kristin
4	Find a document on stagecraft and lighting?	sceneteknikk lyssetting	stagecraft lightning
5	Find a document on cyberbullying?	digital mobbing	Cyberbullying
6	Who is the author of the books about Albert and Skybert?	Albert Skybert	Albert Skybert
7	Find a document about Norwegian recipients of the Nobel Peace Prize?	norske vinnere Nobels fredspris	Norwegian recipients Nobel Peace Prize
8	Find a book on Knut Hamsun and Nazism?	Knut Hamsun Nazisme	Knut Hamsun Nazism
9	Find a document about women in Algeria?	kvinner Algerie	women Algeria
10	Find a play written by William Shakespeare?	William Shakespeare	William Shakespeare

Table 12: Predefined search tasks EXP-III

3.5 Analysis

The analysis software SMI BeGaze version 3.2 was used in all three experiments for exporting eye data and keyboard input, and for creating scanpaths. All screen recordings were analysed in BeGaze. IBM SPSS Statistics 22 was applied for statistical analyses.

3.5.1 EXP-I – Dataset

The eye tracking data from experiment 1 was assessed for quality by looking closely at the raw data. The original dataset consisted of 2184 trials. A total of 19 trials were excluded as follows; two because participants had accidently pushed the space bar before solving the task, four due to data loss and 13 incorrect target responses (for instance, nine participants mistook a calculator for a phone). For each trial that was removed, the corresponding trial for comparison conditions were excluded for both that person and the matching participant. The final dataset comprised of 2056 tasks, 514 for each stimulus type and equally distributed across the four conditions and the two participant groups.

3.5.2 EXP-II & EXP-III – Dataset

Each participant completed ten search tasks each for both the experiments on query formulations. The total data set for these experiments consisted of 400 searches for each of the search systems in EXP-II and EXP-III, equally distributed across the two groups.

3.6 Summary

Selecting suitable participants where observed behaviour may be used to generalise the potential needs of a larger population is both a difficult and an important part of experimental design. Recruiting representative students with dyslexia is challenging, because researchers have no access to lists of potential candidates where subjects can be randomly selected for participation. This privacy protection provides an important ethical framework for research involving disabilities.

Recruitment for this study was based on volunteering participants, which has the disadvantage of potentially introducing a bias. Further, according to Callens, Tops & Brysbaert (2012), dyslexic students may have coping skills above average compared to other dyslexics and possibly less comorbidity. Further, although it is reported that an increasing number of dyslexics are attending higher education (Tops et al., 2012), more severe dyslexics are probably underrepresented in this study, since they probably chose not to attend higher education.

Announcements were made in several fora and students from different fields of study were included to get a representative sample of a large student population. The exception was the exclusion of library- and information science students, since their especially trained search skills could potentially affect the results.

A between subjects design was applied to compare the behaviour of dyslexics with non-dyslexics. This approach is in accordance with other research involving dyslexics. According to Mortimore and Crozier (2006), such a comparison is essential because it would be difficult to identify problems caused by dyslexia without comparisons with non-dyslexic peers. In the context of universal design, where the aim is to improve accessibility for all, it is also relevant to ensure that measures to improve accessibility for dyslexics are not at the expense of non-dyslexic users. The control group was put together to match the dyslexic students as closely as possible, where the presence or absence of dyslexia was the only controlled variable.

It was assumed that language difficulties or reduced vision could negatively affect the results. A set of inclusion criteria was therefore defined to avoid such sources for errors, namely that the participant's had to be native Norwegian speakers and normal to corrected vision was required. It has been claimed that students may not be aware of their dyslexia (Mortimore & Crozier, 2006). Consequently, a commonly applied screening test confirmed group affiliation for both dyslexics and controls. Applying this test also counteracted the need to view sensitive documents about the participants' dyslexia diagnosis.

Three additional cognitive tests were applied, and the results matched the dyslexia profile described in the research literature regarding reduced short term memory capacity (Smith-Spark & Fisk, 2007), concentration difficulties (Mortimore & Crozier, 2006) and impaired naming skills (Lervåg & Hulme, 2009). Several of these tests are also applied by other researchers involving dyslexic participants (for instance Jeffries & Everatt, 2004; Helland & Asbjørnsen, 2004; Smith-Spark & Fisk, 2007; Solheim & Uppstad, 2011; Dahle & Knivsberg, 2014).

The tests applied in this study are legally obtained, publicly available and easily replicable. This combination of tests may therefore be suitable for other research projects targeted at persons with dyslexia and possibly also other cognitive disabilities, but may also be used for studies involving the general population where measures for memory and attention are requested. Consequently, the test battery applied here may be relatively robust and have transfer value to several other types of studies.

The experimental design using verbal instructions accounted for frequent reading errors reported in the dyslexia group (Snowling, 2000). This seems to be important in experimental design involving users with reading difficulties, but may also be useful if illiterate users are involved in other types of studies. The stimulus applied in the three experiments are publicly available, and may thus be easily used to replicate this study or applied to study other aspects of information searching for all types of user groups. The experimental design applied here may therefore have some transfer value.

The search systems applied in EXP-II and EXP-III have different characteristics, such as different tolerance levels for errors and apply help functions differently. These systems therefore seem sufficient to answer H_2 and H_3, which address the use of query buildings aids and how spelling errors affects search behaviour and performance. Differences in performance in these two systems are presented in Chapter 4. To address the hypothesis regarding including graphic content in result lists, another type of experimental stimuli was needed, since no system was available to answer this hypothesis.

Chapter 4: Results

The main purpose behind this study was to acquire empirical knowledge of how query-formulation was affected by dyslexia and to investigate whether the inclusion of graphic content in result lists could be beneficial for dyslexic users without reducing the usability for non-dyslexics. Three main hypotheses were defined to answer this question. The first hypothesis (H_1) suggests that graphic content in result lists may improve search performance among dyslexic users. The second hypothesis (H_2) predicts that dyslexic users would rely more on query-building aids than non-dyslexics. The third hypothesis (H_3) states that the effect of dyslexia on search performance varies according to the system's fault tolerance.

A total of three experiments were completed to answer the hypotheses. The visual search experiment in EXP-I was related to H_1, while the search tasks in EXP-II and EXP-III addressed both H_2 and H_3. In this chapter, the results from the three experiments are presented and discussed according to the main hypotheses.

Chapter 4 is structured as follows: Results from EXP-I regarding visual search and result list content are first reported and discussed in Section 4.1, before the answer to H_1 is summarised and discussed. This is followed by Section 4.2, which addresses the use of query-building aids, and presents results from both the Google and Bibsys experiment (EXP-II and EXP-III). Section 4.3 follows the same structure as 4.2, but discusses how search performance relates to the tolerance level of the system. Section 4.4 includes other findings, which do not directly relate to the main hypotheses, but have implications for search system design and are therefore regarded as relevant in this context, Section 4.5 provides a summary of the findings from the three experiments.

Experimental results are also presented and discussed in four papers (Berget & Sandnes 2015a; Berget & Sandnes, 2015c; Berget, Mulvey & Sandnes, 2016; Berget & Sandnes, n.d.), see Appendices A-D. A set of sub hypotheses were defined for each experiment and included in the papers. In this chapter, the hypotheses are referred to using the same numbers as the experiment and the hypothesis number in the papers, but with a lowercase h to clearly distinguish the sub hypotheses from the experiments (for instance h_{1-1} for the first hypothesis in Experiment-I) from the main hypotheses in this study (H_1).

4.1 H_1 – Result list design

The first hypothesis concerns result lists, and suggests that adding graphic content to text can improve search performance among dyslexics. H_1 was explored through a visual search experiment (EXP-I) described in Section 3.4.1. Participants located targets in four conditions; one-modality tasks with icons or words in a grid-layout, and dual-modality tasks with icons and words in a grid-layout and a list-layout. The findings are presented and discussed in this section, and also in Berget and Sandnes (2015a) and Berget, Mulvey & Sandnes (2016), see appendices A and B.

For the visual search experiment, a set of four more detailed hypotheses were formulated. First, it was assumed that dyslexics would display longer search times than controls when targets included text (h_1-$_1$). The justification for this hypothesis was the general reading and word recognition difficulties reported among people with dyslexia (Snowling, 2000). For the same reason, it was expected that dyslexics would locate targets more easily in tasks displaying graphic content than verbal content only (h_1-$_2$). The third hypothesis (h_1-$_3$) predicts that dyslexics have a preference for icons over words when presented with both modalities simultaneously. Finally, a hypothesis addressing the efficiency of navigation in visual search tasks was formulated. Based on reduced word recognition and word decoding skills (Snowling, 2000), concentration skills (Mortimore & Crozier, 2006) and reduced short-term memory capacity (Smith-Spark & Fisk, 2007), it was expected that dyslexics would navigate less effectively and experience a higher mental load in visual search tasks generally, as compared to controls (h_1-$_4$).

4.1.1 One-modality displays

Participants were given tasks containing only one modality to explore potential differences in performance during icon search compared to word search. In terms of search times, dyslexics displayed no significant differences in tasks containing graphics only, but took significantly longer than controls on textual tasks (see Table 13). This finding was in accordance with $h_{1\text{-}1}$, in that dyslexics would exhibit longer search times than controls in tasks containing text. These results indicate that replacing words with graphic content will not make any difference for dyslexic users, but be significantly worse for non-dyslexics. Consequently, there are no advantages with replacing words with icons in one-modality displays.

	Dyslexics		**Controls**		
Search times (s)	***M***	***SD***	***M***	***SD***	***p (t-test)***
ICON	5.3	1.0	4.9	1.1	>.05
WORD	5.1	2.0	4.1	0.8	<.04

Table 13: Between subjects search times (s) in one-modality conditions

It was predicted that dyslexics would locate targets more easily in tasks displaying graphic content compared to only textual content ($H_{1\text{-}2}$). In the context of one-modality displays, this hypothesis was contradicted, since dyslexics did not take significantly less time solving the ICON tasks compared to WORD tasks (see Table 14). In contrast, non-dyslexics were faster in textual displays compared to searching icons (see Table 14). This finding may be explained by claims that processing icons require more neurological effort than text (Huang, 2012). The results are in accordance with Benbasat and Todd (1993), who found no general advantages with icons compared to text-based representations among general users, but contradict Niemelä and Saarinen (2000), who reported improved search times for icons.

	ICON		WORD		
Search times (s)	***M***	***SD***	***M***	***SD***	***p (t-test)***
Dyslexics	5.3	1.0	5.1	2.0	>.05
Controls	4.9	1.1	4.1	0.8	<.02

Table 14: Within subjects search times (s) in one-modality conditions

There were no significant within subjects differences regarding fixation durations, number of fixations or scanpath lengths (detailed results for these tests are presented in Berget, Mulvey & Sandnes, 2016, p. 9 and 11). Sireteanu et al. (2008) concluded that children with dyslexia were impaired in serial tasks, but found improvements with age. All participants in EXP-I were adults. The findings by Sireteanu et al. (2008) may thus contribute to explain the similarities in search performance between the two groups in tasks with no letter-like stimulus.

Overall, dyslexics had longer fixation durations and dwell times than controls, on both ICON and WORD tasks, but shorter saccade amplitudes only on WORD tasks (see Table 15, 16 and 17). Rayner (2009) has suggested that longer fixations in search tasks may relate to a higher mental load. Consequently, searching in general may be more demanding for dyslexics than controls, which is in accordance with h_{1-4} and confirms the need to consider dyslexic users when designing search user interfaces. However, simply *replacing* text with graphic content in result lists may not be expedient for either dyslexics or non-dyslexics.

	Dyslexics		Controls		
Fixation durations (s)	***M***	***SD***	***M***	***SD***	***p (t-test)***
ICON	238.6	36.2	217.3	29.7	<.05
WORD	241.5	36.5	220.4	32.6	=.05

Table 15: Between subjects fixation durations in one-modality conditions

	Dyslexics		Controls		
Saccadic amplitude in visual degrees	***M***	***SD***	***M***	***SD***	***p (t-test)***
ICON	5.5°	0.6°	5.7°	0.6°	>.05
WORD	5.1°	0.6°	5.6°	1.0°	<.05

Table 16: Between subjects saccadic amplitude in one-modality conditions

	Dyslexics		Controls		
Dwell total dwell time per item (ms)	***M***	***SD***	***M***	***SD***	***p (t-test)***
ICON	378.6	70.9	317.1	57.8	<.005
WORD	388.8	101.6	317.3	46.1	<.008

Table 17: Between subjects dwell times in one-modality conditions

4.1.2 Dual-modality displays

The function of the dual-modality tasks was to investigate whether the combination of icons and words may support users during result list assessment, which is also the topic of Berget, Mulvey and Sandnes (2016), appendix B. According to Jerde et al. (2011), increased set-sizes may negatively affect cognitive workload and performance measures, such as response times. Search performance among dyslexics has also been reported to decrease as a function of the number of distractors (Moores et al., 2011). Consequently, dyslexics could potentially take longer locating items in dual-modality displays compared to searching one modality. The implications of effect-size are important when considering adding a modality to a user interface, because the accessibility will not improve if increased set-sizes have a negative effect.

Dyslexics displayed similar search times in all the tasks with a grid-layout (see Table 18). Increasing set-sizes by combining icons and words in the same layout did therefore not negatively affect search times. This is in agreement with Vidyasagar and Pammer (1999), who found no differences in search times among dyslexics with set-sizes of 10, 24 or 36 items, but discovered a significant difference in displays with 70 items.

	Dyslexics	
Search times (s)	***M***	***SD***
ICON	5.3	1.0
WORD	5.1	2.0
ICON/WORD-GRID	5.3	1.5

Table 18: Search times (s) among dyslexics in conditions with grid-layout

Dyslexics took significantly longer than controls on both the dual-modality tasks, which again confirmed h1-1, in that dyslexics would take longer than non-dyslexics when targets contained text. Both groups solved tasks faster in the dual-modality list-layout compared to the three other conditions (see Table 19 for within subjects search times for dual-modality displays).

	ICON/WORD-GRID		**ICON/WORD-LIST**		
Search time (s)	***M***	***SD***	***M***	***SD***	***p (t-test)***
Dyslexics	5.3	1.5	4.4	1.0	<.01
Controls	4.3	0.8	3.7	0.8	<.01

Table 19: Within subjects search times (s) in dual-modality conditions

Both groups had longer mean fixation durations in the list-layout compared to the other modalities (see Table 20 for dual-modality conditions). According to Rayner (2009), increased fixation durations suggest a higher mental load. Nevertheless, the search performance was improved when items in one modality were closer together, and switching from one content type to another required a saccade outside the parafoveal range (to see the objects sharply enough to process them). These results imply that adding a modality may be useful for dyslexics, and that the second modality may be used to confirm a target when it is found, which is discussed further in Section 4.1.4.

Fixation duration (ms)	ICON/WORD-GRID *M*	*SD*	ICON/WORD-LIST *M*	*SD*	*p (t-test)*
Dyslexics	219.6	24.0	284.5	42.3	<.001
Controls	209.8	31.1	281.2	48.1	<.001

Table 20: Within subjects fixation durations (ms) in dual-modality conditions

Dyslexics had on average more fixations in both the dual-modality displays compared to controls (see Table 21), but no significant differences in fixation durations. (Detailed results of measures that were not found to be significantly different may be found in Berget, Mulvey & Sandnes, 2016, page 8-12). The lacking group differences in either of these displays may imply a generally higher mental load for both participant groups, but that presenting in list format with sufficient spacing between each content type may remove the distracting effect which is added by a second modality.

Number of fixations	Dyslexics *M*	*SD*	Controls *M*	*SD*	*p (t-test)*
ICON/WORD-GRID	20.5	5.6	17.5	3.1	<.04
ICON/WORD-LIST	13.6	3.0	11.5	1.9	<.01

Table 21: Between subjects number of fixations in dual-modality conditions

It was predicted that dyslexics would more easily locate targets in tasks displaying graphic content than verbal content only (h_{1-2}), which was partly confirmed. Both dyslexics and controls solved the dual-modality list condition faster than all the remaining conditions. Consequently, it seems likely that including words (or icons) may have a supportive effect for dyslexics (and non-dyslexics) over icons (or words) alone. However, this finding only seem to apply if the text information is outside central vision when searching icon lists, and vice versa.

4.1.3 User preferences

The results from the visual search experiment were also analysed according to user *preferences* when icons and words were displayed simultaneously. According to Nielsen and Levy (1994), users do not always prefer the conditions in which they perform the best. This view is supported in the dyslexia research, for instance by Williams and Hennig (2015a). According to Friedman and Bryen (2007) and Santana et al. (2012), a preference for graphic content is included in a majority of accessibility guidelines targeted at dyslexic users. Santana et al. (2012) specifically includes icons in their review of existing guidelines. However, contradicting results are reported regarding the function of graphic content in text based material.

Houts et al. (2006) concluded that pictures may increase comprehension among dyslexics if they are closely linked to the text, and also emphasised the importance of images with few distracting details. The icons and words applied in the tasks represented exactly the same concepts. Graphic and textual content were therefore assumed to be closely linked. Moreover, the icons had relatively few details, and no colours. Consequently, the graphic content should be consistent with Houts et al.'s (2006) specifications. Nevertheless, icons did not seem useful when presented with words in a grid-layout, where both content types were within central vision concurrently.

In contrast, studies such as Beacham and Alty (2006) and Brante et al. (2013) concluded that pictures may distract attention away from the text, and consequently not support reading comprehension among dyslexics. However, these studies report findings from learning material and text comprehension, and the results may not be directly transferable to this experiment, where the graphic content had a slightly different function.

It was assumed that dyslexics would prefer graphic content over text (h_{1-4}) due to decoding difficulties (Snowling, 2000). Huang (2012) has suggested that icon processing in general requires more cognitive effort than processing text. Consequently, non-dyslexic users were expected to prefer words over icons.

Areas of interest (AOIs) were defined for each icon and word. In the dual-modality displays, mean dwell times were calculated per type (icon and word). In the grid-layout, where both content types were within the range of central vision concurrently, there were no preferences for any content type in any of the user groups. The only significant difference was found in ICON/WORD-LIST, where dyslexics spent a greater part of the time fixating on icons than controls (see Table 22). Regarding within subjects differences, the only significant difference was that dyslexics spent more time looking at icons than words in the ICON/WORD-LIST condition, *p*<.02 (see Table 22 for descriptive statistics of mean dwell times).

Mean ∑ AOI dwell time in ICON/WORD list (s)	**Dyslexics** *M*	*SD*	**Controls** *M*	*SD*	***p (t-test)***
Icon AOIs	2.3	0.8	1.7	1.0	<.03
Word AOIs	1.4	1.0	1.6	0.9	>.05

Table 22: Group differences in AOI dwell times

Ratio dwell time icon vs. word	**ICON/WORD-GRID** *M*	*SD*	**ICON/WORD-LIST** *M*	*SD*	***p (t-test)***
Dyslexics	3.1	3.6	1.1	0.6	<.016

Table 23: Within subjects ratio between icon and word fixations

Ratios expressing the relationship in viewing time on icons versus words (dwell time icons / dwell time words) were also calculated (see Table 23). Some users navigated both content types, while others fixated on only one modality (see Figure 31). Dyslexics had significant within subjects differences with higher ratios in ICON/WORD-LIST than ICON/WORD-GRID. Consequently, the dyslexic participants spent a higher portion of the dwell time on icons than words in the list-layout compared to the grid-layout which implies that they preferred icons to words in the list-layout. Controls did not prefer words over icons in any of the dual-modality tasks. Based on these results, h_{1-3} which assumed that dyslexics would prefer icons over words was considered to be partly supported. However, the differences in the two layouts imply that arrangement of content has a moderating effect on these preferences.

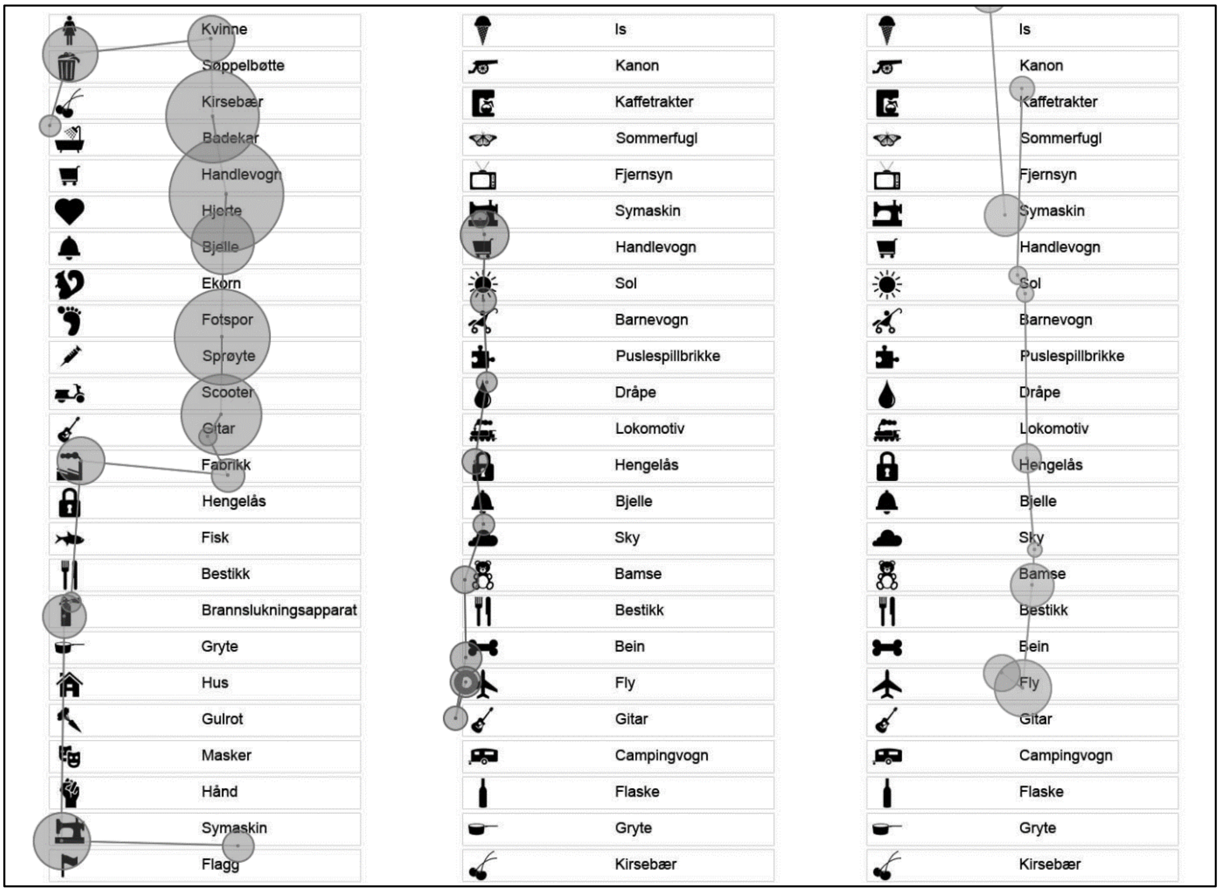


Figure 31: Fixations on both content types (left), icons only (middle) and words only (right)

4.1.4 Content-switching

Content-switching is here referred to as a saccade from one content type to another to verify a target (see Figure 32). This strategy was applied in a majority of searches by both groups (*M*=77.5%, *SD*=25.2% for dyslexics and *M*=61%, *SD*=30.1% for controls). No significant differences were found in the use of this strategy, p>.05.

According to Huang (2012), icons may be more cognitively demanding to process than text. Correct decoding of icons may also depend on characteristics such as gender or age (see Berget & Sandnes, 2015b, appendix E, for differences found during the screening of icons during piloting for EXP-I). During content-switching from icons to words, textual content may assist users in understanding the icons or confirming correct interpretation. In contrast, graphic content may be useful and support word decoding for users who struggle with reading difficulties. Consequently, it was assumed that content-switching from icon to word would be equally applied by both user groups, while movements from words to icons would be more common in the dyslexia group.

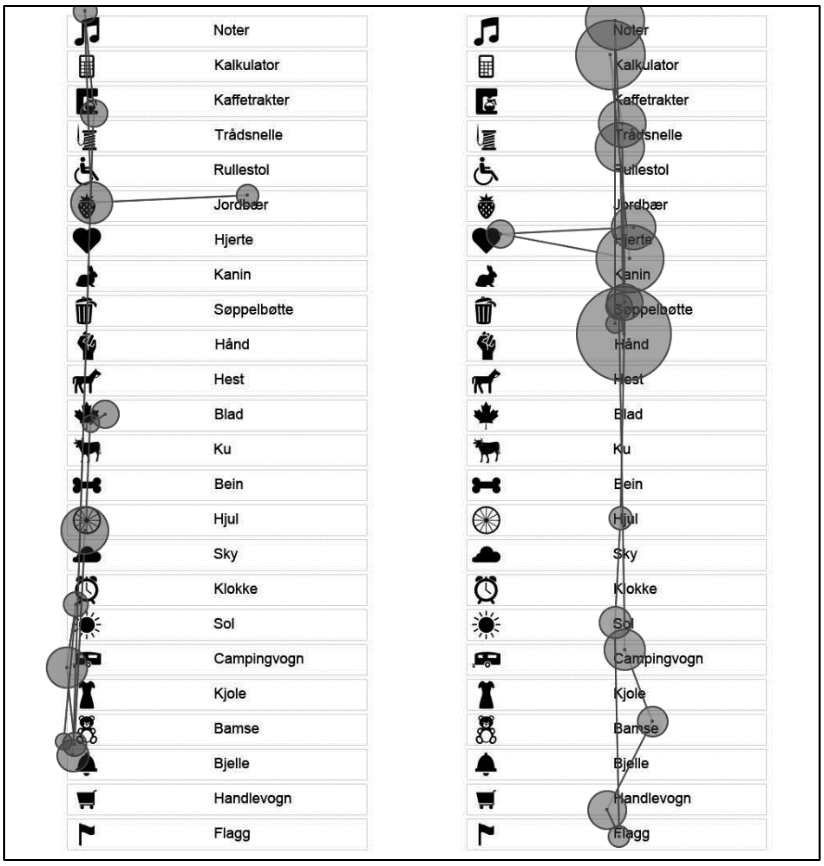


Figure 32: Examples of content-switching towards word (left) and towards icon (right)

Both groups had the highest portion of content-switching from icons to words, which is consistent with findings that decoding icons may be cognitively demanding (Huang, 2012). The only significant difference was the higher portion of content-switching towards words compared to icons within the dyslexia group (see Table 24). However, this behaviour may relate to a higher overall usage of icons in the list condition.

Portion of content-switching (%)	Towards word *M*	*SD*	Towards icon *M*	*SD*	*p (t-test)*
Dyslexics	79.1	21.8	20.9	21.8	<.001

Table 24: Direction of content-switching for dyslexics in %

Several participants commented on the presence of icons in the list-layout as distracting, because they felt obliged to double-check the target, although they were sure about the decoding of the word. This statement also applied to saccades from icons to words. However, since the tasks in the list-layout produced the fastest search-times, this may probably not have been an essential problem. Nevertheless, it seems like a plausible conclusion that in list-layouts users may prefer to navigate mainly one content type, and then apply the other for verification purposes. In these cases, dual-modalities may support search.

4.1.5 H_1 – Result

H_1 predicts that «*graphic content in result lists improves search performance among dyslexic users*». Dyslexics showed no significant improvements in search performance during icon versus word search, while controls did significantly worse in tasks including icons only compared to words. Consequently, H_1 was contradicted in one-modality displays.

In dual-modalities, graphic content seemed more useful for dyslexics than in one-modality displays, a finding that supports H_1. However, it seems probable that for dual-modalities to be useful, the two content types should not be presented within central vision concurrently. This finding suggests that increasing the set-sizes in result lists may not necessarily reduce search performance among dyslexics. Further, the effect of adding a modality seems to also be closely related to layout, and affects both dyslexic and non-dyslexic users.

4.1.6 H_1 – Further thoughts

The conclusion to H_1 is that dyslexic users may benefit from the inclusion of icons in result lists. These findings may have transfer value to other user interface design, for instance menus. However, the experimental results suggest that in addition to including two modalities, layout also matters. Most result lists in search systems applied today present results in list-layouts. Consequently, the inclusion of icons in these lists may be preferable by dyslexic participants. In contrast, in interfaces with array-based layouts, the addition of icons may not be useful. Although there does not seem to be any preferences among controls for either icons or words in list-layouts, it would most likely not reduce the usability for non-dyslexic users to add a modality.

It may be regarded as a positive outcome that there were no contradicting results regarding the two user groups, which according to Newell and Gregor (2000) is a potential obstacle towards universal design. With an aim of universally design search user interfaces, it would not be proper to introduce new graphic interface designs that improved the accessibility for one group at the expense of others.

A possible limitation regarding these results is that the displays were less crowded than a typical result list, which often includes long sentences and surrounding content. This issue may be explored further in future research by replacing single words with longer texts. The increased use of icons by dyslexics in the list-layout may alternatively reflect a preference for left columns compared to right columns, a behaviour which may be connected to typical reading patterns in Western cultures. Another possible limitation of this study is therefore the lack of a fifth condition, with reversed order for icons and words in the list-layout. Such tasks would have rendered possible to study the effect of content order. However, adding another task type would have significantly increased total experiment time, and a fatigue-related effect could have occurred since dyslexics are reported to be more easily fatigued than non-dyslexics (Sireteanu et al., 2008).

One of the sub-hypotheses in EXP-I predicted that dyslexics would navigate less effectively and experience more mental load than non-dyslexics in visual search tasks generally (h_{1-4}). The results from EXP-I partly supports this hypothesis, but it was also found that differences in performance can be modified or even removed depending on the layout and the modality of information. These findings suggest that the negative effect of dyslexia during visual search may be counteracted by well-designed user interfaces, which is in compliance with the universal design mode of thought. However, both the modalities of information and the layout (among others the placement of the two modalities and the spacing between different modalities) must be accounted for during the design process. Consequently, advice on including graphic content in Web accessibility guidelines should probably be succeeded by some details concerning how to present two modalities concurrently for increased accessibility for both dyslexic and non-dyslexic users.

This experiment did not take into account users with other disabilities, such as visual impairments. Further experiments need to be carried out to ensure that the inclusion of graphic content does not negatively affect the accessibility for other users, which would counteract the idea behind universal design.

4.2 H_2 – Query-building aids

According to Gwizdka (2010), query formulation is the phase of the search process which requires the highest cognitive load for all users. In addition to identifying proper search terms, these words must be named and spelled correctly. Snowling (2000) emphasised impaired spelling skills as one of the most common markers for dyslexia, and Catts et al. (2002) reported reduced naming skills among dyslexics. Consequently, it was hypothesised in H_2 that query formulation will be most difficult for dyslexic users, and that query-building aids are utilised extensively by these participants.

Google provides two query buildings aids that may be suitable during query formulation: an autocomplete function and a red line displayed beneath assumedly misspelled words. Bibsys Ask does not offer any such help functions, and could therefore not contribute to answer H_2. However, the participants were allowed to use external resources if they did not manage to solve a task in Bibsys Ask during EXP-III. Except for one direct search in a dictionary, all the external queries were inputted in Google. Findings from these external searches in EXP-II are therefore discussed in this section as well. Detailed results are also presented in two papers (Berget & Sandnes, n.d.; Berget & Sandnes, 2015c, see Appendices C and D).

4.2.1 Autocomplete as query-building aid

Findings from the Google experiment (EXP-II) revealed that dyslexics did not rely more on the autocomplete function than controls (see Table 25), which contradicted H_2. The result was unexpected, since this function was assumed to benefit dyslexic users the most. A general reluctance to use query refinement tools is reported in the research literature (Dennis et al., 1998). This reluctance is, among others, explained by the cognitive load associated with selecting the proper query suggestion, which may in certain cases be spelled quite similar to the surrounding suggestions. The findings from EXP-II may thus be in accordance with these findings. However, another explanation may be found in the eye-data, which revealed that dyslexic users both fixated significantly less on the screen and on the autocomplete list during query input than controls (see Table 25).

	Dyslexics		**Controls**		
Search behaviour	***M***	***SD***	***M***	***SD***	***p (t-test)***
Use of autocomplete function per task	0.33	0.18	0.33	0.19	>.05
Portion of fixation time at screen while typing	73%	27%	90%	11%	<.02
Portion of queries with fixations on autocomplete	57%	25%	71%	20%	<.05

Table 25: Fixation times and autocomplete usage, SD denote standard deviations

The autocomplete function was utilised the same amount by both groups, which implies that such a function may be equally useful for both dyslexics and non-dyslexics. Based on Zipf's law (see Section 2.4.3), it is reasonable to assume that autocomplete should be utilised more frequently, since this function should reduce the workload related to query input. However, the high tolerance for spelling errors in Google also reduces the demands for correctly spelled queries, since users do not have to invest much time in formulating correct queries. Nevertheless, this search strategy may increase the risk of irrelevant result lists, since Google may not always apply the proper spelling correction, as discussed in Sections 4.4.4 and 4.4.5.

4.2.2 Autocomplete as dictionary

The use of the autocomplete function was essentially different in EXP-III from the behaviour observed in EXP-II. In EXP-III, there was a widespread use of the autocomplete function when the students used Google as a support tool for solving tasks in Bibsys Ask. In these cases, the autocomplete function was looked at and used in almost all of the searches, and seemed to be utilised as a dictionary.

Many dyslexics commented that they usually started all searching in Google to find the proper spelling of words before pasting whole queries into the search interface in the library catalogue. Such a strategy may reduce the mental load and effort required to compile the proper query syntax in Bibsys Ask. This approach was observed during tasks with difficult spelling of search terms, where students in both groups searched Google directly on the first query before even trying to solve the task in the library catalogue. Several students explained this behaviour as a result of past experiences with faulty queries in Bibsys Ask. If the students had not been allowed to apply this strategy, it is likely that the search performance in the library catalogue would have been heavily affected.

The strategy of utilising Google as a dictionary was found in both participant groups, and may imply that when looking for the correct spelling of a term, the autocomplete is regarded as useful. However, when searching for information directly in Google, the high tolerance for spelling errors seem to reduce the need for using the autocomplete. This is in accordance with Ward et al. (2012), who concluded that autocomplete functions are mostly applied for spelling correction, and that such functions are more useful in exploratory tasks and at the beginning of the search process.

Queries inputted in Google were in many cases less formal and longer than the queries submitted to Bibsys Ask. For instance, one dyslexic student inputted the query «*nobelsfredspris*» (correct query *Nobels fredspris*, the Nobel Peace Prize) in the library catalogue, a query that returned no results. The student then inputted a query in Google's search interface: «*Norske vinnere av nob*» (Norwegian winners of Nob), whereas a correct query was suggested by Google's autocomplete. This query was copied and pasted into Bibsys Ask. The increased query lengths in Google compared to Bibsys Ask may be a response to a high tolerance for misspellings or that the auto-complete function suggests longer queries.

Participants who seemed uncertain about the spelling of terms typically inputted a few letters only in Google before looking at or selecting queries from the auto-complete list. For instance, one dyslexic user inputted «*nob*» on the task on Norwegian Nobel Peace Prize winners and «*knut*» on the Knut Hamsun and Nazism task. These findings suggest that autocomplete may be helpful when users are uncertain about the spelling of words. However, in contrast to Bibsys Ask, which is a bibliographic database, Google offers full text search. Longer informal queries may thus be more purposeful in Google than Bibsys, which may also explain increased query lengths.

During EXP-III, a total of 19 dyslexics and 13 controls searched Google after submitting an erroneous query in Bibsys Ask, instead of modifying the query in the library catalogue. However, many participants actually resolved the misspellings themselves when inputting the query in Google. In these cases, the students may have solved the tasks in Bibsys Ask without using Google. This finding may imply that users rather go directly to a more tolerant and familiar system if errors occur than spend much time troubleshooting. However, such behaviour can also be caused by a lack of confidence, which according to Jordan, McGladdery and Dyer (2014) is often found among dyslexics.

Some participants reported that they normally used Google more extensively before starting searching the library catalogue, and that they commonly found the correct query in Google and then copied and pasted it into the library catalogue search field. However, in the experimental setting they wanted to try to solve the tasks in Bibsys Ask as instructed. Consequently, the use of external resources may be even more common than reported in this study. Several dyslexics reported that they commonly used Google instead of actual online dictionaries for spell checking of Norwegian words. Others used Google translate for English terms. This behaviour may indicate that users regard Google as more accessible compared to online dictionaries.

4.2.3 Notifications of misspellings

Another query-building aid which was utilised in the search engine was the red line beneath misspelled words. Dyslexic users corrected errors more frequently than controls when the red line appeared, but the usage of this function was not prevailing. Although such a function may be useful, it can be distracting for the users if too many query-building aids are available at the same time.

Changing spelling was the most frequent problem solving approach within both groups (other problem solving approaches are discussed in Section 4.4.2). Dyslexic appeared to be especially aware of common mistakes such as double consonants and compound words, and rephrased queries accordingly, for instance:

> $Q_{1\text{-}1}$: «*scene teknikk*», $Q_{1\text{-}2}$: «*sceneteknikk*».

Another dyslexic applied different ways of spelling, based on double consonants, in addition to the nd- and dt-sound which are also problematic for many dyslexics:

> $Q_{2\text{-}1}$: «*Sigrid Undsett*», $Q_{2\text{-}2}$: «*Sigrid Unsett*», $Q_{2\text{-}3}$: «*Sigrid Unnsett*», $Q_{2\text{-}4}$: «*Sigrid Unnsedt*» (...) (correct spelling Sigrid Undset)

In one of the tasks where Google was applied as an external resource, a relevant result list was presented despite a misspelled query. However, the student did not realise the error, and assumed it was correct since a result list was returned. Consequently, the student continued searching Bibsys Ask with the faulty query and took more than three minutes to solve the task. This behaviour supports the general principle of user interface design that users rely on feedback (Norman, 2010). In this case the red line in the query was not sufficient for the user to discover the error, and it might have been beneficial if Google made the user more clearly aware of the error.

4.2.4 H_2 – Result

H_2 predicts that «*Dyslexic users rely more on query-building aids than non-dyslexics*». The results from the Google experiment (EXP-II) contradicts this hypothesis, since dyslexic users did not rely more on help functions such as autocomplete compared to controls. However, in EXP-III, it was more common for dyslexic users than controls to utilise Google when struggling to solve tasks in Bibsys Ask, which supported H_2. Consequently, H_2 was only partly confirmed, and it seems to depend on the context whether dyslexics rely more on query-building aids than non-dyslexics. The results from EXP-II and EXP-III suggest that query building aids may be useful, but that in systems with a high tolerance for errors such functions seem less important than in systems with a low tolerance for errors. The fault tolerance is discussed further in Section 4.3.

The autocomplete function was applied differently in the two experiments. When solving tasks directly in Google, students did not rely much on query building aids. The eye data revealed that in many searches the students did not even look at the query suggestions, and several students did not look at the screen at all during query input. The latter was especially evident in the dyslexia group, and may be related to a lack of keyboard skills or reduced short term memory capacity caused by the dyslexia. However, when Google was applied as a dictionary for faulty queries in the library catalogue, the autocomplete list was much more actively used, in many cases Google was applied in the same manner as a dictionary. This is in accordance with Ward et al. (2012), who found that autocomplete functions are mainly used for spelling correction.

4.2.5 H_2 – Further thoughts

The implications of the findings in EXP-II and EXP-III seems to be that autocomplete may be useful in certain contexts for some users. However, the intense focus on the keyboard during query input suggests that there may be a need for better keyboard skills among the end users. According to Nicolson and Fawcett (1990), some dyslexics find it difficult to use keyboards, among others due to reduced motor skills and coordination. The finding that dyslexics fixate less on the screen while inputting terms also implies a need to re-think how to best support users during query formulation, and is an example of how eye tracking can be a useful method to document search behaviour.

Another option is to revise of the autocomplete function. White and Marchionini (2007) have suggested that autocomplete functions should be displayed in real-time, before the first results are retrieved, rather than in retrospect for users to notice them, which is the case in Google. However, White and Marchionini (2007) claim that real-time suggestions may also affect the users negatively, and cause incorrect search paths. Query suggestions displayed during input entails that suggestions change while the user inputs text. For instance, if a user search for Bob Dylan in Google and does not fixate on the screen while inputting text, the query suggestion «*bob dylan*» (see Figure 33) will not be noticed when the words «*bob d*» are inputted in the search box. However, if the user misspells the next letter, assuming the spelling is «*bob dilan*», the autocomplete list will no longer be helpful (see Figure 34). In such cases autocomplete functions are not supporting query formulation.

Figure 33: Query suggestions for «*bob d*». Google and the Google logo are registered trademarks of Google Inc., used with permission

Figure 34: Query suggestion for «*bob di*». Google and the Google logo are registered trademarks of Google Inc., used with permission

Figure 35: Query suggestion for «*bob dil*». Google and the Google logo are registered trademarks of Google Inc., used with permission

In this example, the correct suggestion will reappear when the user has entered the next letter, and a red line is then displayed beneath the misspelled last name (see Figure 35). This function can make the user aware of the misspelling. Nevertheless, if the user does complete the search with the query «*bob dilan*», results for Bob Dylan will still be returned due to the spelling correction implemented in Google. Such examples contribute to the conclusion that spelling correction seems more helpful than autocomplete for users who do not fixate on the screen during query input. Fault tolerance is discussed further in Section 4.3.

Findings from the experiments indicate that dyslexia will cause erroneous queries both in systems with query building aids and in systems that offer no such help functions. However, the autocomplete function and the red line beneath misspelled queries seemed to counteract several errors, since the submitted number of errors was lower than the inputted errors in the Google searches. This finding indicates that query building aids may assist dyslexic users during query formulation, which is a positive finding for dyslexic users. The results also indicate that libraries may need to rethink their search user interface designs. In Norway this seems necessary to accommodate the forthcoming juridical deadlines for universally designed ICT solutions. However, it also seems essential for such changes to be able to compete with widely popular Web search engines such as Google.

4.3 H_3 – Fault tolerance

Misspellings are commonly associated with dyslexia (Snowling, 2000). Consequently, dyslexics were expected to make more misspellings during query formulation than non-dyslexic users. However, H_3 suggested that the negative effect of dyslexia that would result in faulty queries would be counteracted by a high tolerance for spelling errors. EXP-II applied a highly tolerant system, while EXP-III used a system with no tolerance for misspellings. Both these experiments are presented in this section, presenting the results from the systems with a high tolerance for errors first, followed by findings from the system with a low tolerance. Detailed findings are also included in two papers (Berget & Sandnes, n.d; Berget & Sandnes, 2015c, see Appendices C and D).

4.3.1 High tolerance

EXP-II addressed the effect of dyslexia in a system with a high tolerance for errors that was assumed to counteract the effect of dyslexia. In this experiment dyslexics were expected to take approximately the same time solving a task compared to controls (h_{2-1}) and submit the same number of queries as non-dyslexics (h_{2-2}). It was also hypothesised that dyslexics would submit longer queries than controls, to increase the precision of result lists, and thus reduce the work load associated with result list assessment (h_{2-3}). Finally, it was predicted that dyslexics would make more spelling errors than controls (h_{2-5}).

EXP-II revealed no significant differences in the search performance of dyslexic users compared to controls, except for the amount of misspellings (see Table 26). Dyslexic users did not take significantly longer solving tasks or submit more queries per task compared to controls (see Table 26), which confirmed h_{2-1} and h_{2-2}. However, they did not submit longer queries, which contradicted h_{2-3}. Only the amount of inputted and submitted spelling errors were significantly different (see Table 26), confirming h_{2-5}.

	Dyslexics		**Controls**		
Query summary	***M***	***SD***	***M***	***SD***	***p (t-test)***
Search time (s) per task	54.5	17.3	45.0	14.8	>.05
Number of queries per task	1.6	0.3	1.4	0.3	>.05
Number of inputted terms per query	3.4	1.1	2.9	0.5	>.05
Number of submitted terms per query	3.5	1.1	3.0	0.5	>.05
Inputted spelling errors per query	0.4	0.3	0.1	0.1	<.001
Submitted spelling errors per query	0.3	0.3	0.1	0.1	<.003

Table 26: Query summary, SD denote standard deviations

The higher amount of misspellings did not have any effect on overall search performance. Consequently, dyslexics and controls performed equally well searching Google, and it seems likely that a high tolerance for errors counteracted the effects of dyslexia on search performance, caused by slow and erroneous query-formulation. This assumption is supported by several examples from the search sessions, for instance the query «*Nygsjeringe Nild forfatter*» (correct query *Nysgjerrige Nils forfatter*), where the dyslexic user got relevant results despite four misspellings in one query and did therefore not have to resubmit a corrected query.

4.3.2 Low tolerance

In contrast to the findings in the Google experiment, it was expected that dyslexia would have a more significant effect on search performance in Bibsys Ask. This system has no query-building aids and no tolerance for spelling errors. In the bibliographic database it was predicted that spelling difficulties, slow writing speed and reduced rapid naming skills would cause longer search times in the dyslexia group compared to controls (h_{3-1}). It was also expected that dyslexics would formulate more queries per task (h_{3-2}), both due to the lack of search aids and because misspelled queries would cause more query reformulations. Moreover, it was presumed that dyslexics would include as few words as possible in each query due to the high demands for correct spelling, to counteract erroneous queries (h_{3-3}). Finally, it was expected that dyslexics would make more misspellings than the controls (h_{3-4}).

In the search system with a low tolerance for errors, dyslexics took significantly longer solving each task (see Table 27) compared to controls. This finding was in accordance with $h_{3\text{-}1}$ and MacFarlane et al. (2010). Increased time usage may be explained by a higher number of queries and more time spent on external Websites. Mortimore and Crozier (2006) have reported that dyslexics may struggle to concentrate on a task for a long time. Consequently, a fatigue-related effect could have caused increased search times. However, time usage did not increase during the experiments, which implies that more accessible search user interfaces may potentially reduce search times for all the tasks.

	Dyslexics		Controls		
Query summary	***M***	***SD***	***M***	***SD***	***p (t-test)***
Search time (s) per task	47.8	14.6	29.0	8.2	<.001
Number of queries per task, overall	2.5	0.6	1.7	0.3	<.001
Number of queries per task, Bibsys	2.1	0.5	1.6	0.3	<.001
Number of queries per task, external	0.4	0.3	0.1	0.1	<.001
Portion of misspelled queries, Bibsys	44%	10%	26%	14%	<.001

Table 27: Query summary, SD denote standard deviations

Dyslexics have been reported to display slower reading speeds than expected according to their age (Hatcher, Snowling & Griffits, 2002). It has also been reported that people with dyslexia may be limited in their choice of words (Connelly, Campbell, MacLean & Barnes, 2006) and have reduced naming skills (Smith-Spark & Fisk, 2007). All these factors may cause more time needed to formulate queries. However, dyslexics did not take significantly longer per query, which was unexpected. However, several search terms, such as author names were included in the verbal instructions, which reduced the need for rapid naming skills to quickly identify search terms. Further, dyslexics submitted more short queries than controls, and time usage may thus have been higher if query lengths had been similar between the two groups throughout the experiment.

Participants in the dyslexia group submitted a significantly higher number of queries overall, in Bibsys and in external resources (see Table 27), confirming $h_{3\text{-}2}$. The experimental findings indicated that the higher number of queries primarily was caused by insufficient or too extensive result lists and misspellings, that led to query modifications or use of external Websites. These results suggest that dyslexia may have a negative impact on search behaviour and performance in systems with a low tolerance for misspellings.

	Dyslexics		Controls		
Query lengths	***M***	***SD***	***M***	***SD***	***p (t-test)***
Submitted terms per query, overall	2.05	0.23	2.25	0.18	<.007
Submitted terms per first query	2.07	0.28	2.16	0.27	>.05
Submitted terms per last query	2.18	0.23	2.30	0.21	>.05
Minimum number of terms per query	1.75	0.33	2.01	0.20	<.006
Maximum number of terms per query	2.35	0.16	2.33	0.22	>.05

Table 28: Query lenghts, SD denote standard deviations

Users with dyslexia also submitted significantly shorter queries than controls (see Table 28), which supported $h_{3\text{-}3}$. However, the first and last queries for each task had similar query lengths to control users. Consequently, the shorter query lengths in the dyslexia group appeared mainly in the intermediate queries. The minimum number of terms per query was significantly lower in the dyslexia group compared to controls; maximum query lengths were similar. This finding indicates that dyslexics apply query reduction when trying to find the correct spelling of terms, for instance:

$Q_{3\text{-}1}$: «*knut hamsun nasismen*», $Q_{3\text{-}2}$: «*knut hamsun*», $Q_{3\text{-}3}$: «*knut hamsun nazismen*».

In the example above, one word was removed in $Q_{3\text{-}2}$, to verify the spelling of the author name, before the subject term was corrected. Query Q_4 shows a similar approach, but in this search each intermediate query contained one term, before both terms were combined in the final query:

$Q_{4\text{-}1}$: «*Sigrid Unseth*», $Q_{4\text{-}2}$: «*Unset*», $Q_{4\text{-}3}$: «*Kristin*», $Q_{4\text{-}4}$: «*Kristin Unset*».

The examples above illustrate how a higher number of queries may be needed to solve a task in a system with a low tolerance for misspellings and no query-building aids. The same search behaviour would not have been required in a more tolerant system, which was documented in EXP-II and discussed in Section 4.3.1. For instance, one student conducted a search in Google after failing to formulate a useful query in Bibsys Ask, and received relevant results in Google despite three misspellings in the query: «*helerystninger i hedemark og oppland*» (correct query helleristninger i hedmark og oppland). The only faulty queries in Bibsys that provided results were queries including the authors Hamsun or Shakespeare, because these terms were mapped to an authority file that directed the user to the correct spelling. These tasks were solved more easily than expected, since the students received results despite errors.

Norman (2002) claims that feedback is a basic principle in user design. However, feedback lacks in Bibsys Ask, where the user is left with a cul-de-sac when misspellings occur. If spelling correction is not implemented, a minimum system requirement should be to provide the users with feedback of what part of the query that has no match in the database, for instance «*no results in the database for the term Unseth*» as a response to $Q_{4\text{-}1}$. Such a function could counteract the need for $Q_{4\text{-}2}$ and $Q_{4\text{-}3}$, and consequently cut the number of queries by a half. Feedback of this type would at least provide the user with information on where an error may have occurred, so the user knows where to begin correcting the query.

Dyslexic users made more misspellings in Bibsys Ask compared to controls (see Table 27). The number of errors increased according to the number of search terms, which was also expected, since a higher number of terms most likely increases the total probability of errors to occur. The portion of misspelled queries was significantly higher in the dyslexia group (see Table 27), where 44% of the queries contained spelling errors, compared to 26% in the control group. This finding confirmed $h_{3\text{-}4}$.

Although controls had fewer misspellings, more than a quarter of the queries included mistakes. This result is in accordance with Cucerzan and Brill's (2004) findings that users generally make a significant amount of misspelling in queries. However, the number of errors in EXP-III was higher than the findings of Cucerzan and Brill (2004), who reported 10-15% erroneous queries.

The amount of errors related to query lengths and the use of difficult or unfamiliar words, such as *helleristninger* and *Algerie*. Previous research has reported that dyslexic adults may memorise the spelling of familiar words, but have more difficulties with orthographic patterns in unfamiliar words, which may result in a higher number of errors (Kemp, Parrila and Kirby, 2009). The search logs from the Bibsys Ask experiment support this finding.

The findings from the Bibsys Ask experiment indicate that dyslexia has a pronounced negative effect on information retrieval in search systems with no tolerance for errors and causes longer search times and a higher number of queries needed to solve a task.

4.3.3 H_3 – Result

H_3 stated that «*The effect of dyslexia on search performance relates to the system's tolerance for errors*». The search system with no tolerance for errors caused longer search times and more queries among dyslexics. Consequently, the presence of dyslexia seemed to have a strong impact on overall search performance in system with a low tolerance for errors. The same results were not found in the system with a high tolerance for errors. This finding confirms H_3.

4.3.4 H_3 – Further thoughts

The results from EXP-III indicates that dyslexics may be impaired in information searching, and that search performance is heavily affected by dyslexia. However, the results from the Google experiment (EXP-II) are quite positive regarding dyslexics, suggesting that the negative impact dyslexia has on information search may be counteracted by well-designed search user interfaces including a high fault tolerance. Further, since many non-dyslexics also make spelling errors during query formulation, alterations in search user interfaces that benefit dyslexics may also benefit people without dyslexia. This study is an example where empirical user knowledge may be used to alter user interfaces to become more universally designed. More adequate design may probably both counteract a disability, and allow users to utilise ICT on the same level as people without such an impairment.

The results from the Bibsys Ask experiment (EXP-III) are disappointing from a library perspective, in that the search systems seems inaccessible or less accessible to both students with dyslexia and students in general. However, since many systems already provide a high tolerance level for errors, it seems likely that little effort is actually needed to make library catalogues more accessible, and that altering the search user interface is probably a good investment.

The dyslexic students found searching search user interfaces with low fault tolerance frustrating and time consuming, and therefore preferred search engines such as Google, also for academic content. However, students without dyslexia also made errors in this system and expressed similar discontent towards the library catalogue. It seems likely that alterations which accommodate dyslexics may also improve the usability for non-dyslexics, a finding which is consistent with other research on universal design (McCarty & Swierenga, 2010). Consequently, implementing a high tolerance for errors seems to be a useful measure to develop more universally designed search user interfaces for a broader user group than dyslexics. Making more accessible search systems also seems vital if libraries should be able to compete with commercial search engines such as Google.

4.4 Other findings

The results presented above directly answers the main hypotheses. However, during the search experiments, certain other behaviour was observed that may affect search user interface design as well. These results are presented below, and includes converting information needs into queries, coping strategies among dyslexics, query language, extensive misspellings, misinterpretations, Cul-de-sacs and the need for library support.

4.4.1 Converting information needs into queries

In the Google experiment participants with dyslexia sometimes entered queries that may imply difficulties in transferring information needs into precise queries. Dyslexics often have impaired naming skills (Catts et al., 2002), which may cause difficulties in finding the proper query terms.

Several users replicated the exact wording or parts of the of the task instructions, which resulted in queries consisting of both relevant and irrelevant search terms. For instance, one dyslexic inputted the query: «*hvor stor er blinken på liggende skyting i skisyting*» (what is the size of the target in prone position in biathlon, last word misspelled). In comparison, users in the control group typically submitted queries containing only the significant terms, for instance: «*skiskyting liggende blink*» (biathlon prone position target). This finding may indicate that dyslexics find it more difficult to convert information needs into concise queries. Although queries close to everyday-language may not be problematic in full-text databases, such strategies will not be applicable in bibliographic databases. However, more research is needed to investigate this issue further.

4.4.2 Coping strategies

According to Borkowski, Estrada, Milstead and Hale (1989), dyslexics often develop coping strategies to compensate for their impairment. Students with dyslexia were therefore expected to apply a broader range of problem solving strategies ($h_{3\text{-}5}$) than controls. However, the only strategy applied significantly more frequent in the dyslexia group was the use of external Websites (see Table 29). Google was the only external Website applied, except for one direct search in an online dictionary. This behaviour implies that users may prefer the search user interface of Google compared to Bibsys Ask. The finding is also in accordance with the results from the pre-interviews (see Section 3.2.3), where students in general expressed a much more positive attitude towards Google compared to Bibsys Ask.

	Dyslexics		Controls		
Problem solving strategies	***M***	***SD***	***M***	***SD***	***p (t-test)***
Portion of queries with changed spelling	46%	13%	40%	20%	>.05
Portion of queries in external Web sites	27%	11%	16%	15%	<.017
Portion of queries with increased query lengths	16%	8%	27%	24%	>.05
Portion of queries with decreased query lengths	11%	9%	10%	12%	>.05
Portion of queries with replacement of words	6%	9%	7%	16%	>.05

Table 29: Query lenghts, SD denote standard deviations

An evident strategy for modifying an erroneous query is to change the spelling of at least one term. Other possible approaches are for instance to include, replace or exclude difficult words. These strategies were especially obvious in the task to find the author of the books about Albert and Skybert (in Norwegian the Sk is pronounced sh, making it potentially difficult to spell). When solving this task, 25% of the dyslexics formulated queries including «*Albert Åberg*». The last name of this children's book character (Åberg) was not included in the task. Consequently, the query was based on prior knowledge and included in the queries, thereby avoiding a difficult term. Such a strategy was also applied in 25% of the searches by the non-dyslexics. Word replacement might thus be a common strategy for all types of users, and is not necessarily a characteristic marker of dyslexic users.

Strategies involving word replacements were also applied in intermediate queries. For example, one dyslexic student used this approach when searching for «*Helleristninger i Hedmark og Oppland*» and Hedmark was regarded difficult to spell ($Q_{5\text{-}4}$):

> $Q_{5\text{-}1}$: «*helleristinger*», $Q_{5\text{-}2}$: «*helleristinger hedmark*», $Q_{5\text{-}3}$: «*helleristinger hedemark*», $Q_{5\text{-}4}$: «*helleristinger i oppland*»

The replacement strategy was found in many of the queries inputted by dyslexic users, trying to avoid difficult terms. This finding is in accordance with Riddick (1996), who found that a common strategy among dyslexics is to avoid words which are difficult to spell or to ask for help from others if avoidance is not an option. However, replacement approaches may not always be successful in bibliographic databases such as Bibsys Ask, since the indexed content is more limited than in full text databases.

In bibliographic databases, queries must match metadata terms such as names, titles or subject headings. Exceptions to this demand is the use of synonym controls or thesauri, where non-preferred terms may be mapped to the content of the database. In these cases, a replacement approach may be productive. Search systems should therefore as a minimum support such functions to accommodate users who exhibit this type of query modification strategy.

4.4.3 Query language

According to White and Marchionini (2007), autocomplete functions may cause incorrect search paths. Such search behaviour was also observed in the Google experiment. For instance, in some searches the participants unwarily selected an English query, because they entered terms with identical spelling in Norwegian and English (for instance «*museum Egypt*»). English is reported to be challenging for many Norwegian dyslexics (Helland & Kaasa, 2005). Consequently, result lists in English may not be beneficial for dyslexics, except for when English Web pages are accepted in the result list.

Several students mentioned the presence of English query suggestions as a distraction in the autocomplete list in Google. There was no language option for the autocomplete function. Consequently, the user is offered query suggestions in different languages simultaneously. For instance, the list in Figure 36 contains queries in Norwegian, Swedish and Danish. Further, the first suggestion in Figure 36 returns results in all the three languages. In such cases, it would have been helpful if the users could easily filter results by language. Google has such an option in the advanced search, and sometimes at the top of the results list, but this could also be displayed on the simple search page for easier access.

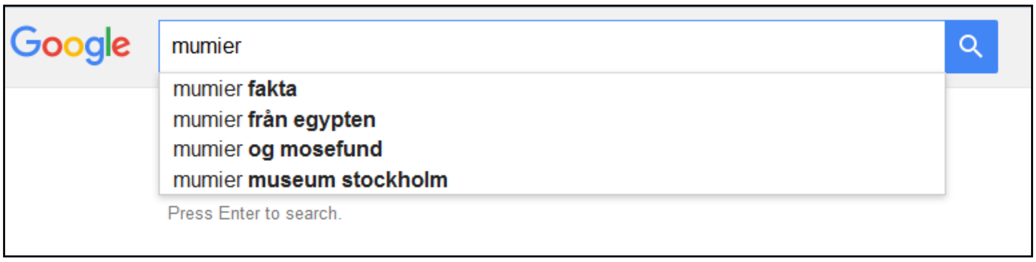


Figure 36: Query suggestions in three different languages. Google and the Google logo are registered trademarks of Google Inc., used with permission.

4.4.4 Extensive misspellings

In some of the queries conducted by dyslexics in EXP-II the autocomplete was not helpful because of extensive misspellings. For instance, one dyslexic student submitted the query «*shamaika*» to solve the task about Jamaica. However, the search system did not recognise what the user intended to search for, and suggested «*shamaila*»: The student selected this option and searched with this query for some time before realising the error. In this search, autocomplete did not facilitate the search, possibly due to comprehensive misspellings. This finding may indicate a need for spell checkers that better accommodate spelling errors of dyslexics, which according to Moats (1996) may differ from spelling mistakes of other users.

4.4.5 Misinterpretations

According to the findings of the two experiments on information retrieval, a high tolerance for errors may be preferable for both user groups, but has a larger impact on the search performance of dyslexics than controls. However, the higher tolerance for errors does not always enhance the search process, since the system may misinterpret queries. Figure 37 shows the result list from the Google search «*ole brumm sitat undre*» (translates Winnie the Pooh quote wonder), which is a correct query when searching for Winnie the Pooh quotes in Norwegian about wondering or thinking. However, Google interprets the last term as an error, and returns results of *under* (translates under or below) instead of *undre* (wonder).

In the example in Figure 37, Google gives the impression that the search results are related to the query «*ole brumm sitat undre*», since the query is repeated below the images. However, a closer look at the actual results shows that the word *under* is emphasized in each document excerpt, and that the word *undre* is not included. In this case, Google has altered the search query and presented a significantly different result list from what the user expected. In such cases, Google may return documents that do not make sense for the users, who may end up spending much time in the result lists trying to find a suitable document to fulfil the information need. For users with dyslexia, it may be even more difficult to see the misinterpretation, since the words *undre* and *under* only differs in the order of the two last letters.

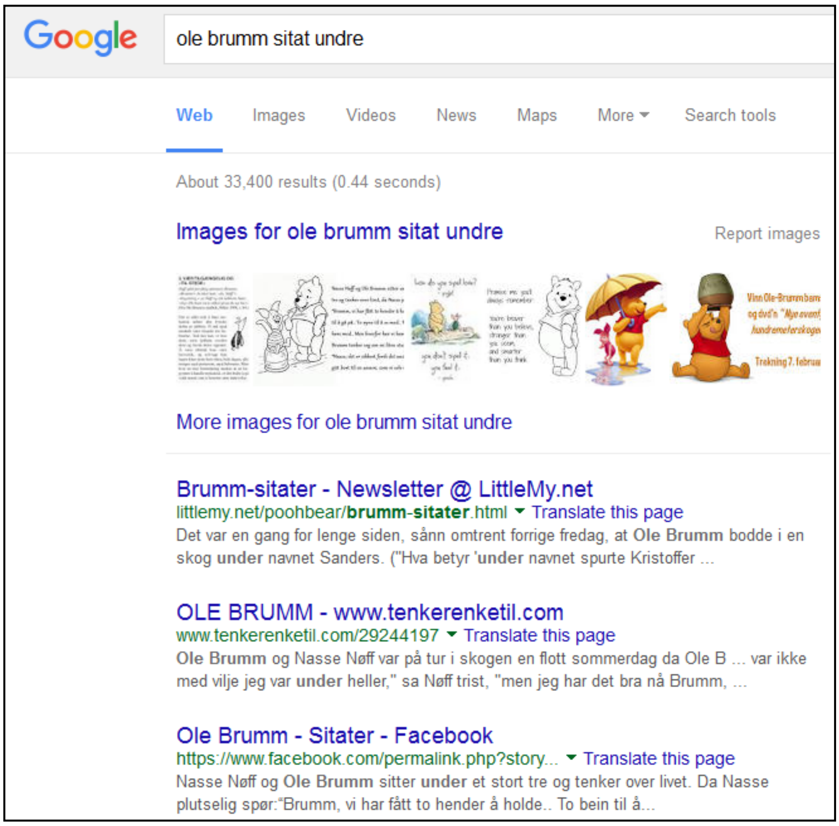


Figure 37: Result list showing documents which does not directly relate to the query. Google and the Google logo are registered trademarks of Google Inc., used with permission

Another potential challenge with automatic correction can occur in cases where the search system may struggle to understand what the user actually is searching for. For instance, the Norwegian words *kornpris* (the price on grain) and *kronprins* (Crown Prince) include almost the same letters. If a user inputs the query «*kornprins norge*», it is difficult for the system to predict whether the user meant to search for grain prices in Norway or the Crown Prince of Norway. In such cases, the system should either provide the user with two options and allow the user to select the correct one, or at least try to analyse the results based on previous searches. For instance, if the previous search included words such as the Royal Family or wheat, it may be easier to try to predict what the user actually meant.

In EXP-II there were some cases where the system did not recognise what the user was searching for. For instance, while solving the task on alcohol consumption and boating, a dyslexic user entered «*alkohol båtkjørinmg*» (alcohol boating [the latter spelled incorrectly]) and then selected «*alkohol bilkjøring*» (alcohol car driving) in the autocomplete list. Such errors may be caused by decoding errors, which are commonly reported among dyslexic users (Shaywitz & Shaywitz, 2005). However, it may also be a result of incorrect use of autocomplete, where the users uncritically trust the suggestions by the search system. This is in accordance with the findings by White and Marchionini (2007).

A great trust in Google's suggestions was also observed during the Bibsys Ask experiment when two dyslexic students searched Google after failing to formulate a useful query in the library catalogue. In both cases, the participants searched for Norwegian winners of the Nobel Peace prize, and got the suggestion «n*orske nobel vinnere*» (Norwegian nobel winners) by Google. This query was copied and pasted into Bibsys Ask. In theory this was not an incorrect suggestion, since the two words have meaning separately. Nevertheless, in this particular task, the two last terms had to be written as one compound word: «*norske nobelvinnere*» (Norwegian winners of Nobelprize). The participants did not notice this error, and submitted three additional queries each to the library catalogue before realising there might be something wrong with the query.

4.4.6 Cul-de-sacs

In one of the Google tasks a dyslexic user got no results when trying to solve the rear-facing car seat task with the query «*alder seteintruksjon*» (age seat instruction, last word misspelled). Google did not provide any suggestions, but advised the user to verify the spelling, try different keywords, more general keywords or fewer keywords. The same advice was generally suggested by Bibsys Ask when no results were retrieved. The students expressed uncertainty about how to proceed when this happened in Google, since this rarely occurs and was therefore not expected. Consequently, the participant struggled for more than 4.5 minutes to solve the task. This finding suggest that users may become too dependent on help functions in certain systems, and that anxiety may arise when these functions do not prove helpful.

4.4.7 Library support

A majority of dyslexic students mentioned asking a librarian for help as a commonly used approach. One of the dyslexic participants had given up searching the library catalogue altogether, and never tried searching for literature because it was too time-consuming and usually ended in failure. This statement is in accordance with the claims by Mortimore and Crozier's (2006), that a majority of dyslexic students in higher education wanted or used extra library support. However, it is in contrast to Helland (2002) and Katopol (2005) who claimed that dyslexic students were afraid to approach the library, and therefore not used the library services as much as they would have wanted.

4.5 Summary

The main conclusion from the visual search experiment was that graphic content may be useful in result lists. However, the usefulness of adding a modality depends on the layout and presentation, namely that the two modalities should not be presented within central vision concurrently. Based on the query formulation experiments, it seems likely that a high tolerance for errors may counteract the negative effect dyslexia has on information search. Further, fault tolerance seems more important than query-building aids for both dyslexic and non-dyslexic users. However, autocomplete functions may be useful when users are uncertain about the spelling of certain search terms, where the search engine is used in the same manner as a dictionary. A few other findings were also reported from these experiments, which may have implications for the design of both search user interface design and other user interfaces for dyslexics. The main findings from the three experiments that may have implication for search user interface design are summarised in Table 30.

Query formulation
High tolerance for errors
Implement query building aids such as autocomplete function and red line
Allow for full text search
Synonym control
Allow users to turn help features on/off (for instance the red line beneath assumed misspellings)
Easily available language settings for spelling corrections in search box, autocomplete list and result lists
Spell checkers that consider typical dyslexic misspellings
Autocomplete function that considers that users may look away from screen during input
Result lists
Suggest queries if misspellings occur
Sufficient feedback of assumed errors
Sufficient feedback when queries are corrected by the system
Sufficient help when searches fail, no dead ends
Dual-modality result lists where graphic and textual content is not presented within central vision concurrently

Table 30: Issues to consider for search user interface design

Chapter 5: Conclusion

This study started with the aims to better understand how dyslexia affects information search, and to get a better understanding of how to design search user interfaces that are accessible for dyslexics. This closing chapter summarises the answer to the research question and the three main hypothesis in Section 5.1. This is followed by a discussion of the implications of the study in Section 5.2 and the main limitations of this study in Section 5.3. Future research is discussed in Section 5.4 before the chapter is ended by some final remarks in Section 5.5.

5.1 Research question answers

Information retrieval has been reported as difficult for all users (Aula, Khan & Guan, 2010). However, the cognitive profile of dyslexia discussed in Section 2.2 suggests that searching may be even more difficult for dyslexic users, both during query formulation and result list assessment. Consequently, the starting point for this study was the following research question: *How are query-formulation and result list assessment in online search systems affected by dyslexia?* Three hypotheses were formulated to address this issue:

> **H_1**: *Graphic content in result lists improves search performance among dyslexic users*
> **H_2**: *Dyslexic users rely more on query-building aids than non-dyslexics*
> ***H_3**: The effect of dyslexia on search performance relates to the system's tolerance for errors*

Three experiments were conducted to answer these hypotheses, which together contributed to provide an answer to the research question.

The first hypothesis addresses result list content and EXP-I (described in Section 3.4.1) investigated whether the inclusion of graphic content could support dyslexic users during result list assessment (H_1). This hypothesis was confirmed. However, for graphic content to be useful, the two modalities (icons and text) should be displayed with a distance apart so that they are not in central vision concurrently. Another main finding was that differences in search performance during visual search tasks between dyslexics and non-dyslexics can be modified or removed, depending on the layout and modality of information. In other words, the negative effect of dyslexia may be counteracted through appropriate layout and the modalities the information is presented in.

The two remaining hypotheses addresses query formulation. The second hypothesis (H_2) regards the use of query building aids based on the assumption that such help functions would be more utilised by dyslexics, who struggle with spelling (Snowling, 2000). The third hypothesis (H_3) addresses fault tolerance, and it was investigated whether the effect of dyslexia on search performance related to the tolerance level for errors in the system. EXP-II (see Section 3.4.2 and 3.4.3) and EXP-III (see Section 3.4.2 and 3.4.4) were conducted in two different search systems, the Web search engine Google and the library catalogue Bibsys Ask, and results from both of these experiments were used to answer H_2 and H_3.

Regarding query-building aids, the dyslexics did not rely more on functions such as autocomplete compared to controls while solving tasks in Google, among others due to an intense focus on the keyboard during query input. However, in the library catalogue tasks it was more common for dyslexics to turn to Google and utilise query-building aids. In these searches, Google was used almost like they used an online dictionary. Consequently, H_2 was only partly confirmed.

The two query formulation experiments resulted in quite different results. In the system with a low tolerance for errors (EXP-III), dyslexia had a negative impact on query formulation. Dyslexia affected search performance in terms of the number of misspellings, search times and the number of queries needed to solve a search task. However, the same effect was absent in the system with a high fault tolerance (EXP-II), confirming H_3. Consequently, it was concluded that differences in search performance between dyslexics and non-dyslexics can be modified or even removed, depending on the search user interface design.

Based on the answers to the three hypotheses, the answer to the research question is that both query formulation and result list assessment can be negatively affected by dyslexia. Consequently, dyslexics can be impaired in information searching compared to non-dyslexic users. However, a positive finding from all the experiments is that relatively simple measures in search user interface design may reduce or actually remove the negative impact of dyslexia, both during result list assessment and query formulation. For instance, dyslexics do not seem impaired when searching systems with a high fault tolerance. Further, the dual-modality list-layout also counteracted or removed the difficulties dyslexics displayed in visual search tasks. Consequently, in the context of universal design, the conclusion may be reversed, stating that certain search systems are discriminating towards dyslexics, who are not able to perform on the same level as non-dyslexics.

5.2 Implications of this study

Since online search systems were introduced, information searching has changed from being an activity for a few occupational groups to become a common activity for most people. Actually, search systems are offered in most Web sites and applications today, from library catalogues, to search engines and ticket machines. Consequently, search user interfaces and information search behaviour have become central topics in HCI and library and information science research.

Another central development in HCI research is the increasing awareness of and focus on accessible technology and universal design. However, as discussed in Section 2.1, there is still a gap in the knowledge of how cognitive impairments affect the interaction with search user interfaces. The overall purpose of this study was to narrow the gap by contributing with empirical knowledge about how dyslexia affects information searching. Such user knowledge has implications for the development of accessibility guidelines and best practise.

This study focused on search engines and bibliographic databases. However, the findings may be transferred to other search user interfaces containing a search function. There is no reason to assume that query formulation will differ much when searching other databases, except for characteristics related to the search functions in that particular system. Users who are searching for train tickets to a certain station or looking up a word in an online dictionary are probably equally exposed for misspellings as when searching for general information in a Web search engine or trying to locate a book in a bibliographic database. Consequently, the results from these experiments may have transfer value to a variety of settings and search systems.

Query formulations have been discussed at a general level in the library- and information science literature, among others the difficulties in transferring information needs into queries. According to Gwizdka (2010), query formulation is the phase during information retrieval that requires the most cognitive load. A basis for this study was the assumption that query-formulation is even more demanding for dyslexics.

The results from these experiments indicate that dyslexia may heavily affect information retrieval. In this experiment, spelling skills had a high impact on information retrieval. However, another typical characteristic of dyslexia is a reduced short-term memory capacity. MacFarlane et al. (2012) suggested that a reduced short term memory also affects information search, among others during relevance assessment. Consequently, there is a need to include characteristics such as reading and writing skills and short term memory capacity in general information search behaviour models.

According to universal design principles, it is important to accommodate the needs of as many users as possible. Consequently, measures that increase the accessibility for one user group should not be at the expense of others. The main difficulty for dyslexic users is related to reading and writing. Such difficulties are also reported among deaf users (Boldyreff, Burd, Donkin & Marshall, 2001) and non-native speakers. However, other users also frequently make spelling errors, for instance in stressful situations or due to fatigue or reduced concentration (McCarthy & Swierenga, 2010). Consequently, many non-dyslexic users share characteristics with dyslexics, at least in certain situations or contexts. Measures to improve search user interfaces for dyslexics may therefore be useful for many other, non-dyslexic users.

Findings from the information retrieval experiments indicated that although the search performance differed between the two groups, there were also concurrent needs related to commonly occurring errors and the utilisation of query-building aids. Consequently, the search process can probably be enhanced for *both* dyslexics and non-dyslexics through the same measures. This conclusion also applied to the visual search experiment, where the dual-modality list-layout proved most effective for both user groups, although they had slightly different preferences regarding the utilisation of graphic versus textual content.

According to the gap-model two different measures may improve accessibility; changing the environment and strengthening the individual. Universal design may be a useful approach to change the environment. The information retrieval experiments identified a gap between the demands from systems with low tolerance for errors and the search skills of dyslexic users. In certain cases, this gap was also found for non-dyslexic users. The gap may be reduced or even removed through more accessible search systems. This perception was supported by the Google experiment, where the gap was significantly narrower than in the Bibsys experiment.

WCAG (W3C, 2008) is now widely applied as the de facto standard for Web accessibility. However, revised guidelines that address the importance of decreasing demands for correct spelling could probably increase Web accessibility for dyslexic users (and probably many non-dyslexics). This issue is also discussed in Berget, Herstad and Sandnes (2016). Changes in search user interfaces would be a measure referred to in the gap-model as changing the environment. However, the suggestion to develop better keyboard skills among dyslexic users (Berget & Sandnes, 2015c), which could enable more fixations on the screen while inputting text, is an example of strengthening the person's abilities. Consequently, a combination of these measures may narrow the gap and increase accessibility of online search.

In certain cases, it may not be possible to accommodate all users through one design. This issue is addressed in the usability pyramid. However, the findings of this study imply that in the context of dyslexics and information search, there are clear measures in the lowest level of the pyramid that will improve the accessibility. Consequently, it does not seem necessary to develop adaptations particularly for dyslexics. This view is also in accordance with the universal design mentality.

Taylor (1968) emphasised that users may have difficulties expressing an actual information need, and emphasised the role of librarians as intermediates between the user and the system. A consequence of the changed role of the library to a more digital, self-assisted service, is that the role of librarians as such mediators may be reduced or removed. Consequently, the search systems must accommodate users accordingly, and take over the responsibility for helping users with query formulations.

Creating universally designed search user interfaces are important in the context of digital libraries, but also in relation to distant learning students, for users who are reluctant to ask librarians for help and in libraries that are unstaffed during certain hours in the evening. Results from this study suggest that inaccessible search user interfaces in libraries represent a serious obstacle for dyslexic users, but also for non-dyslexics who also struggled searching the library catalogue.

Dyslexic students are found to be especially reluctant to make use of the library (Riddick, Farmer & Sterling, 1997; Helland, 2002). At the same time, there is the paradox that dyslexics depend on more help from librarians to utilise the library services (Mortimore & Crozier, 2006), which was also mentioned by participants in this study. By increasing the accessibility of the search user interfaces, these services may be more easily available for dyslexics. There is also a possibility that an increased sense of coping may reduce the library anxiety reported among dyslexics.

The main goal of this study was to investigate the effect of dyslexia on information retrieval. However, the experiment in the library catalogue also revealed that Norwegian academic library systems are not accessible to dyslexic users. Further, findings from the Google experiment shows that by altering the search user interfaces, academic library services in Norway may become much more accessible for both dyslexics and non-dyslexics. The conclusion that library catalogues are inaccessible for dyslexic users is in accordance with Rutledge's (2002) findings, in that dyslexics were generally disadvantaged or excluded from libraries.

Libraries have a social responsibility to assist all users. In Norway, it is actually embedded in the library jurisdiction that public libraries are obliged to assist *all users* in the country (Kulturdepartementet, 1985). The same should apply to academic libraries, by providing universally designed services that accommodate the needs of *all students*. Inaccessible libraries may have serious consequences for dyslexic students. If dyslexics do not manage to retrieve academic papers from reading lists, or fail to find the appropriate information for a bachelor or master thesis, higher education becomes inaccessible. The consequences of excluding dyslexic students from the library services may therefore actually decrease the possibilities to complete a higher degree, and thus reduce the opportunities in the labour market. Consequently, accessible search is necessary to provide satisfactory library services to dyslexic users and make higher education accessible.

5.3 Limitations

Assistive technology is often discussed in relation to accessibility. Such technology may be useful for dyslexics, for instance screen readers and speech recognition software (Draffan, Evans & Blenkhorn, 2007). Assistive technology was not commonly used by any of the participants in this study. Moreover, the purpose of this study was to better understand how to develop accessible search for dyslexics without the need of such technology, which is an important aspect of universal design. Consequently, assistive technology has not been discussed in this thesis. However, results may have been different for the users if they had applied speech synthesis or spelling correction tools.

As discussed in Section 4.1.6, there are also possible limitations in that the displays applied in the visual search tasks contained short words and were less crowded than regular result lists. Further, there should have been a fifth modality which contained a reversed list with words in the left column and icons in the right column. Such stimulus could have been used to investigate whether the preference for icons may also be related to a preference for left column content (which may reflect typical Western reading patterns).

Certain limitations may be found in the recruitment procedures. One criterion was that the participants had to be students. It seems likely that the users who struggle the most with dyslexia probably do not choose to attend higher education. Consequently, the population may be biased. Another potential weakness was that the study relied on students volunteering to participate. However, it was difficult to avoid this procedure, since no records of students with dyslexia are available (which is a good thing in terms of ethics and the protection of privacy, since information about diagnoses is sensitive information). Further, the channels used to announce the need for participants may also have introduced bias, by for instance reaching more students from certain fields of study.

Finally, experimental settings may cause behaviour that is not in accordance with regular behaviour. For instance, the users may apply other systems to retrieve information than the ones used in this experiment or search for quite different topics. However, since the systems that were used represent quite typical search systems, this is hopefully not a serious problem.

5.4 Future research

New questions emerged during the analysis of the results from the three experiments. For instance, the effect of layout on visual search performance, which is not entirely addressed in this study since reversed lists (words in left column and icons in right column) were not included. Consequently, this should be investigated further.

There is a lack of characteristics related to impairments in existing information searching models. In Wilson's (1999) general model (discussed in Section 2.4.3), psychological and demographic variables are included as intervening variables. This model could be revised to also include individual abilities related to disabilities, both physical and cognitive variables. An advantage of such an extension would be to more clearly display that these are characteristics which should be included in research on information seeking behaviour. A possibility is therefore to investigate further how commonly used theories and models may integrate impairments as potential user characteristics.

Another question regards keyboard design, and whether there are ways to support dyslexics users during text input, so they may focus more on the screen while inputting text. Other topics of interest would be how the severity of dyslexia affects search performance, and whether there are differences regarding the presence of dual-diagnoses.

In a wider context, this study shows a need for more research on users with cognitive disabilities, and how they interact with computers in general and search systems in particular. The findings from this experiment suggest that it may be useful to explore further how different cognitive disabilities affect information search, a topic that has received little attention. The experimental design applied here could be used in a study involving for instance people with ADHD or learning disorders in addition to users within the autism spectrum. In addition to uncover the needs of people with cognitive impairments, such user studies may also have the potential to positively affect the accessibility for non-disabled users.

5.5 Final remarks

After this study was completed, the library catalogue Bibsys Ask was replaced by Bibsys Oria. Although the search interface is changed, among others including suggestions of new searches and ways to sort or refine the search, query-building aids such as autocomplete are still not provided. Further, the tolerance level for errors seems to be unchanged. Consequently, the findings in this study may still apply to Norwegian academic library catalogues, even though the search user interface is revised.

References

Abras, C., Maloney-Krichmar, D., Preece, J. (2004) User-centered design. In Bainbridge, W. *Encyclopedia of Human-Computer Interaction*. Thousand Oaks: Sage Publications.

Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends in Cognitive Sciences, 11*(11), 458-465.

Al-Maskari, A. & Sanderson, M. (2011). The effect of user characteristics on search effectiveness in information retrieval. *Information Processing & Management, 47*(5), 719-729.

Al-Wabil, A., Zaphiris, P. & Wilson, S. (2007). Web navigation for individuals with dyslexia: An exploratory study. In C. Stephanidis (Ed.), *Universal Access in Human Computer Interaction: Coping with Diversity: Lecture Notes in Computer Science, 4th International Conference on Universal Access in Human-Computer Interaction, UAHCI 2007, Held as Part of HCI International 2007, Beijing, China, July 22-27, 2007, Proceedings, Part I* (pp. 593–602). Berlin: Springer.

Alexander-Passe, N. (2006). How dyslexic teenagers cope: An investigation of self-esteem, coping and depression. *Dyslexia, 12*(4), 256-275.

Association of College and Research Libraries (2000). *Information literacy competency standards for higher education*. Chicago: American Library Association.

Aula, A., Khan, R. M. & Guan, Z. (2010). How does search behavior change as search becomes more difficult? In E. Mynatt, G. Fitzpatrick, S. Hudson, K. Edwards & T. Rodden (Eds.), *CHI 2010 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, April 10-15, 2010, Atlanta, Georgia, US* (pp. 35-44). New York: ACM.

Aula, A., Majaranta, P. & Räihä, K.-J. (2005). Eye-tracking reveals the personal styles for search result evaluation. In M. F. Costabile & F. Paternò (Eds.), *Human-Computer Interaction – INTERACT 2005: IFIP TC 13 International Conference, Rome, Italy, September 12-16: Proceedings* (pp. 1058-1061). Berlin: Springer.

Australian Human Rights Commission. (2014). *World Wide Web Access: Disability Discrimination Act Advisory Notes ver. 4.0*. Retrieved 3.1.2016 from http://humanrights.gov.au/world-wide-web-access-disability-discrimination-act-advisory-notes-ver-40-2010#transition

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556-559.

Baddeley, A. (2003). Working memory: Looking back and forward. *Nature Reviews Neuroscience, 4*(10), 829-839.

Baddeley, A. D. & Hitch, G. J. (2008). Working memory. In J. Andrade (Ed.), *Memory: Critical Concepts in Psychology, volume III: Working Memory*. (pp. 125-161). London: Routledge.

Badian, N. A. (1999). Reading disability defined as a discrepancy between listening and reading comprehension: A longitudinal study of stability, gender differences, and prevalence. *Journal of Learning Disabilities, 32*(2), 138-148.

Baeza-Yates, R. & Rello, L. (2011). Estimating dyslexia in the Web. In L. Ferres, M. Vigo & J. Abascal (Eds.), *W4A'11 Proceedings of the International Cross-Disciplinary Conference on Web Accessibility* (article no. 8). New York: ACM.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review, 13*(5), 407-424.

Bayard, S., Erkes, J. & Moroni, C. (2009). *Test du Stroop Victoria – adaptation francophone: Matériel, consignes, procedure de cotation et données normatives*. Gignac: Collège des Psychologues Cliniciens spécialisés en Neuropsychologie du Languedoc Roussillion (CPCN-LR).

Bayard, S., Erkes, J. & Moroni, C. (2011). Victoria Stroop Test: Normative data in a sample group of older people and the study of their clinical applications in the assessment of inhibition in Alzheimer's disease. *Archives of Neuropsychology, 26*(7), 653-661.

Beacham, N. A. & Alty, J. L. (2006). An investigation into the effects that digital media can have on learning outcomes of individuals who have dyslexia. *Computers & Education, 47*(1), 74-93.

Becker, S. I. (2011). Determinants of dwell time in visual search: Similarity or perceptual difficulty?. *PLoS ONE, 6*(3), e17740.

Beins, B. C. (2009). *Research methods: A tool for life*. Boston: Pearson

Belkin, N. J. (1980). Anomalous state of knowledge as a basis for information retrieval. *Canadian Journal of Information Science, 5*, 133-143.

Belkin, N. J. (2000). Helping people find what they don't know. *Communications of the ACM, 43*(8). 58-61.

Benbasat, I. & Todd, P. (1993). An experimental investigation of interface design alternatives: Icon vs. text and direct manipulation vs. menus. *International Journal of Man-Machine Studies, 38*(3), 369-402.

Benjamins, J. S, Hooge, I. T. C., van Elst, J. C., Wertheim, A. H. & Verstraten, F. A. J. (2009). Search time critically depends on irrelevant subset size in visual search. *Vision Research, 49*(3), 398-406.

Berget, G., Herstad, J. & Sandnes, F. E. (2016). Search, read and write: An inquiry into Web accessibility for dyslexics. In H. Petrie, J. Darzentas, T. Walsh, D. Swallow, L. Sandoval, A. Lewis & C. Power (Eds), *Universal Design 2016: Learning from the Past, Designing for the Future* (pp. 450-460). Amsterdam: IOS Press.

Berget, G. & Moseid, T. (2011). Mind the Gap: Mellom oppsøkende bibliotektjenester og det universelt utformede bibliotek [Mind the Gap: Between Outreach library services and the universally designed library]. In R. A. Audunson (Ed.), *Krysspeilinger: Perspektiver på bibliotek- og informasjonsvitenskap (pp.* 185-204). Oslo: ABM-media.

Berget, G., Mulvey, F. & Sandnes, F. E. (2016). Is visual content in textual search interfaces beneficial to dyslexic users?. *International Journal of Human-Computer Studies, 92-93*, 17-29.

Berget, G. & Sandnes, F. E. (2015a). The effect of dyslexia on searching visual and textual content: Are icons really useful?. In M. Antona & C. Stephanidis (Eds.), *Universal Access in Human-Computer Interaction: Access to Learning, Health and Well-being: 9th International Conference, UAHCI 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings, Part III* (pp. 616-625). Cham: Springer.

Berget, G. & Sandnes, F. E. (2015b). On the understandability of public domain icons: Effects of gender and age. In M. Antona & C. Stephanidis (Eds.), *Universal Access in Human-Computer Interaction: Access to Today's Technologies, 9th International Conference, UAHCI 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings, Part I* (pp. 387-396). Cham: Springer.

Berget, G. & Sandnes, F. E. (2015c). Searching databases without query-building-aids: Implications for dyslexic users. *Information Research, 20*(4), article 689.

Berget, G. & Sandnes, F. E. (n.d.). Do autocomplete functions reduce the effect of dyslexia on information searching behavior? The case of Google. *Journal of the American Society for Information Science & Technology*, accepted 2015.

Berners-Lee, T. & Cailliau, R. (1992). World-Wide Web. In *Computing in High Energy Physics 92, Annecy, France, 23-27 September 1992*. Retrieved from http://www.freehep.org/chep92www.pdf.

Bertera, J. H. & Rayner, K. (2000). Eye movements and the span of the effective stimulus in visual search. *Perception & Psychophysics, 62*(3), 576-585.

Beverly, C. A., Bath, P. A. & Barber, R. (2007). Can two established information models explain the information behaviour of visually impaired people seeking health and social care information?. *Journal of Documentation, 63*(1), 9-32.

Beymer, D., Russell, D. & Orton, P. (2008). An eye tracking study of how font size and type influence online reading. In D. England (Ed.) *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction – Volume 2* (pp. 15-18). Swinton: British Computer Society.

Boer-Schellekens, L. d. & Vroomen, J. (2012). Sound can improve visual search in developmental dyslexia. *Experimental Brain Research. 216*(2), 243-248.

Boldyreff, C. Burd, E., Donkin, J. & Marshall, S. (2001). The case for the use of plain English to increase Web accessibility. In S. Tilley (Ed.), *Proceedings of the 3rd International Workshop on Web Site Evolution* (pp. 42-48). Los Alamitos: IEEE.

Borgman, C. L., Hirsh, S. G. & Walter, V. A. (1995). Children's searching behavior on browsing and keyword online catalogs: The Science Library Catalogue Project. *Journal of the American Society for Information Science and Technology, 46*(9), 663-684.

Borkowski, J. G., Estrada, M. T., Milstead, M. & Hale, C. A. (1989). General problem-solving skills: relations between metacognition and strategic processing. *Learning Disability Quarterly, 12*(1), 57-70.

Borlund, P. (2003). The IIR evaluation model: A framework for evaluation of interactive information retrieval systems. *Information Research, 8*(3), article 152.

Brajnik, G. (2009). Validity and reliability of Web Accessibility Guidelines. In S. Harper & A. Barreto (Eds.), ASSETS'09, October 25-28, 2009, Pittsburgh, Pennsylvania, USA (pp. 113-120). New York: ACM.

Brante, E. W., Olander, M. H. & Nyström, M. (2013). Exploring the impact of contrasting cases in text and picture processing. *Journal of Visual Literacy, 23*(2), 15-38.

British Dyslexia Association. (n.d.). *Dyslexia Style Guide*. Retrieved 4.2, 2016, from http://www.bdadyslexia.org.uk/common/ckeditor/filemanager/userfiles/About_Us/policies/Dyslexia_Style_Guide.pdf

Bruneau, D., Sasse, M. A. & McCarthy, J. (2002). The eyes never lie: The use of eye tracking data in HCI research. In *Proceedings of the CHI'02 Workshop on Physiological Computing* (no pagination). New York: ACM.

Buckland, M. K. (2008). Reference library service in the digital environment. *Library and Information Science Research, 30*(2), 81-85.

Buckland, M. & Gey, F. (1994). The relationship between recall and precision. *Journal of the American Society for Information Science, 45*(1), 12-19.

Burgstahler, S. (2002). Distance learning: Universal design, universal access. *Educational Technology Review, 10*(1), 32-61.

Burgstahler, S. & Doe, T. (2004). Disability-related simulations: If, when and how to use them in professional development. *Review of Disability Studies, 1*(2), 8-18.

Buscher, G. Cutrell, E. & Morris, M. R. (2009). What do you see when you're surfing?: Using eye tracking to predict salient regions of Web pages. In D. R. Olsen & R. B. Arthur (Eds.), *CHI 2009 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 21-30). New York: ACM.

Byrne, M. D., Anderson, J. R., Douglass, S. & Matessa, M. (1999). Eye tracking the visual search of click-down menus. In M. G. Williams & M. W. Altom (Eds.), *CHI 99 Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 402-409). New York: ACM.

Bødker, S. (2006). When second wave HCI meets third wave challenges. In A. Mørch, K. Morgan, T. Bratteteig, G. Ghosh & D. Svanæs (Eds.), *NordiCHI 2006 Proceedings of the 4th Nordic conference on Human-computer interaction: Changing roles* (pp. 1-8). New York: ACM.

Callens, M., Tops, W. & Brysbaert, M. (2012). Cognitive profile of students who enter higher education with an indication of dyslexia. *PLoS ONE, 7*(6), e38081.

Cantor, A. (1996). The adaptable approach: A practical guide to planning accessible libraries. *Library Hi Tech, 14*(1), 41-45.

Caroll, J. M. & Iles, J. E. (2006). An assessment of anxiety levels in dyslexic students in higher education. *British Journal of Educational Psychology, 76*(3), 651-662.

Case, D. O. (2005). Principle of least effort. In K. E. Fisher, S. Erdelez & L. E. F. McKechnie (Eds.), *Theories of information behavior* (pp. 289-292). Medford, NJ: ASIST.

Case, D. O. (2007). *Looking for information: A survey on research on information seeking, needs, and behavior*. Bingley: Emerald.

Catts, H. W., Gillispie, M., Leonard, L. B., Kail, R. V. & Miller, C. A. (2002). The role of speed of processing, rapid naming, and phonological awareness in reading achievement. *Journal of Learning Disabilities, 35*(6), 509-524.

Center for Universal Design. (2008). *About UD*. Retrieved 17.12.2015 from https://www.ncsu.edu/ncsu/design/cud/about_ud/about_ud.htm

Chang, P.-Y. & Chang, S.-J. L. (2010). National Taiwan Library services for visually impaired people: A study using Sense-Making approach. *Journal of Educational Media & Library Services, 47*(3), 283-318.

Chatman, E. (1996). The impoverished life-world of outsiders. *Journal of the American Society for Information Science, 47*(3), 193-206.

Ciszek, M. P. (2012). Diversifying diversity: Library services for underrepresented groups. *College & Research Library News, 73*(9), 547-549.

Colenbrader, A. (2008). The historical evolution of visual acuity measurement. *Visual Impairment Research, 10*(2-3), 57-66.

Cowan, N. (2008). What are the differences between long-term, short-term and working memory? *Progress in Brain Research, 169*, 323-338.

Craven, J. & Brophy, P. (2003). *Non-visual access to the digital library: The use of digital library interfaces by blind and visually impaired people*. Manchester: Centre for Research in Library and Information Management.

Crotty, M. (1998). *The foundations of social research: Meaning and perspectives in the research process*. London: Sage.

Cucerzan, S. & Brill, E. (2004). Spelling correction as an iterative process that exploits the knowledge of Web users. In D. Ling & D. Wu (Eds.), *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, EMNLP 2004, A meeting of SIGDAT, a Special Interest Group of the ACL, held in conjunction with ACL 2004, 25-26 July 2004, Barcelona, Spain* (pp. 293-300). Barcelona: ACL.

Cutrell, E. & Guan, Z. (2007). What are you looking for?: An eye tracking study of information usage in Web search. In M. B. Rosson & D. Gilmore (Eds.), *CHI'07 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 407-416). New York: ACM.

Czaja, S. J., Sharit, J., Ownby, R., Roth, D. L. & Nair, S. (2001). Examining age differences in performance of a complex information search and retrieval task. *Psychology and Aging, 16*(4), 564-579.

Dahle, A. E. & Knivsberg, A.-M. (2014). Internalizing, externalizing and attention problems in dyslexia. *Scandinavian Journal of Disability Research, 16*(2), 179-193.

Dampuré, J., Ros, C, Rouet, J.-F. & Vibert, N. (2014). Task-dependent sensitization of perceptual and semantic processing during visual search for words. *Journal of Cognitive Psychology, 26*(5), 530-549.

Daneman, M. & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450-466.

Davies, J. E. (2007). An overview of international research into the library and information needs of visually impaired people. *Librarytrends, 55*(4), 785-795.

Dennis, S., McArthur, R. & Bruza, P. (1998). Searching the World Wide Web made easy?: The cognitive load imposed by query refinement mechanisms. In *Proceedings of the 3rd Australian Document Computing Symposium, Sydney, Australia, August 21 1998*. Retrieved from http://eprints.qut.edu.au/61310/1/XIR-bruza-ADCS98.pdf.

Dervin, B. (2005). What methodology does to theory: Sense-Making as exemplar. In K. E. Fisher, S. Erdelez & L. E. F. McKechnie (Eds.), *Theories of information behavior* (pp. 25-29). Medford, NJ: ASIST.

Disability Rights Commission. (2004). *The Web: Access and inclusion for disabled people: A formal investigation conducted by the Disability Rights Commission*. London: TSO.

Dixon, M. (2007). Comparative study of disabled vs. non-disabled evaluators in user-testing: Dyslexia and first year students learning computer programming. In C. Stephanidis (Ed.), *Universal Access in Human Computer Interaction: Coping With Diversity, 4th International Conference on Universal Access in Human-Computer Interaction, UAHCI 2007, Held as Part of HCi International 2007, Beijing, China, July 22-27, 2007, Proceedings, Part I* (pp 647-656). Berlin Heidelberg: Springer.

Dörk, M., Williamson, C. & Carpendale, S. (2012). Navigating tomorrow's Web: From searching and browsing to visual exploration. *ACM Transactions on the Web, 6*(3), article no. 13.

Draffan, E. A., Evans, D. G. & Blenkhorn, P. (2007). Use of assistive technology by students with dyslexia in post-secondary education. *Disability and Rehabilitation: Assistive Technology, 2*(2), 105-116.

Ehmke, C. & Wilson, S. (2007). Identifying Web usability problems from eye-tracking data. In T. Ormerod & C. Sas (Eds.), *BSC-HCI '07 Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI... but not as we know it – Volume 1* (pp119-128). Swinton: British Computer Socity Swinton.

European Parliament. (2002). *eEurope 2002: Accessibility of Public Web Sites and their Content: European Parliament resolution on the Commission communication eEurope 2002: Accessibility of Public Web Sites and their Content*. Retrieved 14.1.2016 from http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+TA+P5-TA-2002-0325+0+DOC+PDF+V0//EN

Evans, W. (2013). 'I am not a dyslexic person I'm a person with dyslexia': Identity constructions of dyslexia among students in nurse education. *Journal of Advanced Nursing, 70*(2), 360-372.

Evett, L. & 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.

Faccioli, C., Peru, A., Rubini, E. & Tassinari, G. (2008). Poor readers but compelled to read: Stroop effects in developmental dyslexia. *Child Neuropsychology, 14*(3), 277-283.

Facoetti, A., Paganoni, P. & Lorusso, M. L. (2000). The spatial distribution of visual attention in developmental dyslexia. *Experimental Brain Research, 132*(4), 531-538.

Ferris, F. L. & Bailey, I. (1996). Standardizing the measurement of visual acuity for clinical research studies. *Ophthalmology, 103*(1), 181-182.

Findlay, J. M. (2004). Eye scanning and visual search. In: J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision and action: Eye movements and the visual world* (pp. 135-159). New York: Psychology Press.

Findlay, J. M. & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. Oxford: Oxford University Press.

Ford, M., Acosta, A. & Sutcliffe, T. J. (2013). Beyond terminology: The policy impact of a grassroot movement. *Intellectual and Developmental Disabilities, 51*(2), 108-112.

Ford, N., Miller, D. & Moss, N. (2001). The role of individual differences in Internet searching: An empirical study. *Journal of the American Society for Information Science and Technology, 52*(12), 1049-1066.

Francks, C., MacPhie, I. L. & Monaco, A. P. (2002). The genetic basis of dyslexia. *The Lancet Neurology, 1*(8), 483-490.

Friedman, M. G. & Bryen, D. N. (2007). Web accessibility design recommendations for people with cognitive disabilities. *Technology and Disability, 19*(4), 205-212.

Fuchs, C. & Obrist, M. (2010). HCI and society: Towards a typology of universal design principles. *International Journal of Human-Computer Interaction, 26*(6), 638-656.

Gallagher, A., Frith, U. & Snowling, M. J. (2000). Precursors of literacy delay among children at genetic risk of dyslexia. *Journal of Child Psychology and Psychiatry, 41*(2), 203-213.

Germanò, E., Gagliano, A. & Curatolo, P. (2010). Comorbidity of ADHD and dyslexia. *Developmental Neuropsychology, 35*(5), 475-493.

Gibbs, J., Appleton, J. & Appleton, R. (2007). Dyspraxia or developmental coordination disorder?: Unravelling the enigma. *Archives of Disease in Childhood, 92*(6), 534-539.

Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N. & Wichansky, A. M. (2002). Eye tracking in Web search tasks: Design implications. In A. T. Duchowski, R. Vertegaal & J. W. Senders (Eds.), *ETRA ’02 Proceedings of the 2002 symposium on Eye tracking research & applications* (pp. 51-58). New York: ACM.

Goldberg, J. H. & Wichansky, A. M. (2002). Eye tracking in usability evaluation: A practitioner’s guide. In: J. Hyönä, R. Radach & H. Deubel (Eds.): *The Mind’s Eye: Cognitive and Applied Aspects of Eye Movement Research I* (pp. 493-516). Amsterdam: Elsevier Science.

Gossen, T., Hempel, J. & Nürnberger, A. (2013). Find it if you can: Usability case study of search engines for young users. *Personal and Ubiquitous Computing, 17*(8), 1593-1603.

Gould, J. H. & Glencross, D. J. (1990). Do children with a specific reading disability have a general serial-ordering deficit? *Neuropsychologica, 28*(3), 271-278.

Granka, L. A., Joachims, T. & Gay, G. (2004). Eye-tracking analysis of user behavior in WWW search. In M. Sanderson, K. Järvelin, J. Allan & P. Bruza (Eds.), *SIGIR '04 Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 478-479). New York: ACM.

Greene, H. H. & Rayner, K. (2001). Eye movements and familiarity effects in visual search. *Vision Research, 41*(27), 3763-3773.

Gross, M. (2005). The imposed query. In K. E. Fisher, S. Erdelez & L. E. F. McKechnie (Eds.), *Theories of information behavior* (pp. 164-168). Medford, NJ: ASIST.

Grudin, J. (2005). Three faces of human-computer interaction. *IEEE Annals of the History of Computing, 27*(4), 46-62.

Guan, Z. & Cutrell, E. (2007). An eye tracking study of the effect of target rank on Web search. In M. B. Rosson & D. Gilmore (Eds.), *CHI ’07 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 417-420). New York: ACM.

Guron, L. M. (1999). *Wordchains: A word reading test for all ages*. London: GL Assessment.

Gwizdka, J. (2010). Distribution of cognitive load in Web search. *Journal of the American Society for Information Science and Technology, 61*(11), 2167-2168.

Habib, L., Berget, G., Sandnes, F. E., Sanderson, N., Kahn, P., Fagernes, S. & Olcay, A. (2012). Dyslexic students in higher education and virtual learning environments: An exploratory study. *Journal of Computer Assisted Learning, 28*(6), 574-584.

Hariri, N. (2011). Relevance ranking on Google: Are top ranked results really considered more relevant by the users? *Online Information Review, 35*(4), 598-610.

Harper, S. (2007). Is there design-for-all?. *Universal Access in the Information Society, 6*(1), 111-113.

Harrison, S., Tatar, D. & Sengers, P. (2007) The three paradigms of HCI. In Alt. chi. Proceedings of CHI’07. New York: ACM.

Hawke, J. L., Olson, R. K., Willcut, E. G., Wadsworth, S. J. & DeFries, J. C. (2009). Gender ratios for reading difficulties. *Dyslexia, 15*(3), 239-242.

Hearst, M. A. (2009). *Search user interfaces*. Cambridge: Cambridge University Press.

Helland, S. (2002). *Å vere student med diagnosen dysleksi: Empirisk analyse av studiesituasjonen for førskulelærarstudentar med spesifikke lese- og skrivevansker* [To be student with a dyslexia diagnosis: An empirical analysis of the study situation for preschool teacher students with specific reading- and writing disorders]. Oslo: Universitetet i Oslo.

Helland, T. & Asbjørnsen, A. (2004). Digit Span in dyslexia: Variations according to language comprehension and mathematics skills. *Journal of Clinical and Experimental Neuropsychology, 26*(1), 31-42.

Helland, T. & Kaasa, R. (2005). Dyslexia in English as a second language. *Dyslexia, 11*(1), 41-60.

Helland, T., Tjus, T., Hovden, M., Ofte, S., & Heimann, M. (2011). Effects of bottom-up and top-down intervention principles in emergent literacy in children at risk of developmental dyslexia: A longitudinal study. *Journal of Learning Disabilities, 44*(2), 105-122.

Henderson, J. M. (2012). Introduction to “Computational approaches to reading and scene perception”. *Visual Cognition, 20*(4-5), 357-359.

Henry, S. L., Abou-Zahra, S. & Brewer, J. (2014). The role of accessibility in a universal Web. In J. P. Bigham, Y. Borodin & L. Carriço (Eds.), *W4A’14 Proceedings of the 11th Web for All Conference* (article no.17). New York: ACM.

Hepworth, M. (2007). Knowledge of information behaviour and its relevance to the design of people-centred information products and services. *Journal of Documentation, 63*(1), 33–56.

Hersberger, J. (2005). Chatman’s information poverty. In K. E. Fisher, S. Erdelez & L. E. F. McKechnie (Eds.), *Theories of information behavior* (pp. 289-292). Medford, NJ: ASIST.

Hillier, R. (2008). Sylexiad: A typeface for the adult dyslexic reader. *Journal of Writing in Creative Practice, 1*(3), 275-291.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Weijer, J. v.d. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Holone, H. & Herstad, J. (2013). Three tensions in participatory design for inclusion. In W. E. Mackay, S. Brewster & S. Bødker (Eds.), *CHI’13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2903-2906). New York: ACM.

Houts, P. S., Doak, C. C., Doak, L. G. & Loscalzo, M. J. (2006). The role of pictures in improving health communication: A review of research on attention, comprehension, recall and adherence. *Patient Education and Counseling 61*(2), 173-190.

Huang, S.-C. (2012). *Are icons pictures or logographical words? Statistical, behavioral, and neuroimaging measures of semantic interpretations of four types of visual information*. Texas: University of Texas.

Huang, S.-C., Bias, R. G. & Schnyer, D. (2014). How are icons processed by the brain? Neuroimaging measures of four types of visual stimuli used in information systems. *Journal of the American Society for Information Science, 66*(4), 702-720.

Huang, S.-M., Shieh, K.-K. & Chi, C.-F. (2002). Factors affecting the design of computer icons. *International Journal of Industrial Ergonomics, 29*(4), 211-218.

Huang, X., Jing, J., Zou, X.-B., Wang, M.-L., Li, X.-H. & Lin, A.-H. (2008). Eye movements characteristics of Chinese dyslexic children in picture searching. *Chinese Medical Journal, 121*(17), 1617-1621.

Humphrey, N. & Mullins, P. M. (2002). Self-concept and self-esteem in developmental dyslexia. *Journal of Research in Special Education Needs, 2*(2), no pagination.

Husain, R. & Asari, M. A. (2006). From card catalogue to OPACs. *DESIDOC Bulletin of Information Technology, 26*(2), 41-47.

Høien, T. & Lundberg, I. (2012). *Dysleksi: fra teori til praksis* [Dyslexia: from theory to practise]. Oslo: Gyldendal.

Høien, T. & Tønnesen, G. (2008). *Ordkjedetesten* [The Word Chain Test]. Bryne: Logometrica.

Iles, J., Walsh, V. & Richardson, A. (2000). Visual search performance in dyslexia. *Dyslexia, 6*(3), 163-177.

International Dyslexia Association, the. (2016). *Definition of Dyslexia*. Retrieved 13.2.2016 from http://eida.org/definition-of-dyslexia/

Internet Live Stats. (n.d.). *Google search statistics*. Retrieved 6.1.2016 from http://www.internetlivestats.com/google-search-statistics/

Irwin, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. In: J. M. Henderson & F. Ferreira (Eds.), *The Interface of Language, Vision and Action: Eye Movements and the Visual World* (pp. 105-133). New York: Psychology Press.

ISO. (2009). *ISO 8596 Ophthalmic Optics: Visual Acuity Testing – Standard Optotype and its Presentation*. Geneva: ISO.

ISO. (2010). *ISO 9241-210:2010: Ergonomics of human-system interaction: Part 210: Human-centred design for interactive systems.* Geneva: ISO.

ISO. (2012). *ISO/IEC 40500:2012: Information technology – W3C Web Content Accessibility Guidelines (WCAG) 2.0*. Geneva: ISO.

Iwarsson, S. & Ståhl, A. (2003). Accessibility, usability and universal design – positioning and definition of concepts describing person-environment relationships. *Disability and Rehabilitation, 25*(2), 57-66.

Jacob, R. J. K. & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind, 2*(3), 573-605.

Jacob, R. J. K., Leggett, J. J., Myers, B. A. & Pausch, R. (1993). An agenda for human-computer interaction research: Interaction styles and input/output devices. *Behaviour & Information Technology, 12*(2), 69-79.

Jamali, H.R., & Asadi, S. (2010). Google and the scholar: The role of Google in scientists' information-seeking behaviour. *Online Information Review, 34*(2), 282–294.

Jeffries, S. & Everatt, J. (2004). Working memory: Its role in dyslexia and other specific learning difficulties. *Dyslexia, 10*(3), 196-214.

Jerde, T. A., Ikkai, A. & Curtis, C. E. (2011). The search for the neural mechanisms of the set-size effect. *European Journal of Neuroscience, 33*(11), 2028-2034.

Jones, M. W., Branigan, H. P. & Kelly, M. L. (2007). Visual deficits in developmental dyslexia: Relationships between non-linguistic visual tasks and their contribution to components of reading. *Dyslexia, 14*(2), 95-115.

Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329-354.

Kacmar, C. J. & Carey, J. M. (1991). Assessing the usability of icons in user interfaces. *Behaviour & Information Technology, 10*(6), 443-457.

Kane, M. J., Thuholski, S. W., Hambrick, D. Z., Wilhelm, O., Payne, T. W. & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*(2), 189-217.

Kaptelinin, V., Nardi, B., Bødker, S., Carroll, J., Hollan, J., Hutchins, E. & Winograd, T. (2003). Post-cognitivist HCI: Second-wave theories. In G. Cockton & P. Korhonen (Eds.), *CHI EA '03: CHI '03 Extended Abstracts on Human Factors in Computing Systems* (pp. 692-693). New York: ACM.

Katopol, P. (2005). Library anxiety. In K. E. Fisher, S. Erdelez & L. E. F. McKechnie (Eds.), *Theories of information behavior* (pp. 235-238). Medford, NJ: ASIST.

Kessels, R. P. C., Zandvoort, M. J. E. V., Postma, A., Kappelle, L. J. & Haan, E. H. F. D. (2000). The Corsi Block-Tapping Task: Standardization and normative data. *Applied Neuropsychology, 7*(4), 252-258.

Khalid, H. M. (2006). Embracing diversity in user needs for affective design. *Applied Ergonomics, 37*(4), 409-418.

Kim, K.-S. & Allen, B. (2002). Cognitive and task influences on Web searching behaviour. *Journal of the American Society for Information Science & Technology, 53*(2), 109-119.

Kommunal- og Moderniseringsdepartementet. (2013). Forskrift om universell utforming av informasjons- og kommunikasjonsteknologiske løsninger [Regulation on universal design of information and communation technological solutions]. Retrieved 1.3.2016 from https://lovdata.no/dokument/SF/forskrift/2013-06-21-732?q=Forskrift%20om%20universell%20utforming%20av*

Korrekturavdelingen. (n.d.). *Frekvensordliste: De vanligste norske ord* [Word frequency list: The most common Norwegian words]. Retrieved 20.02.2013 from http://www.korrekturavdelingen.no/K4Frekvensordliste500VanligsteNorsk.htm.

Kraft, P. & Bansler, J. P. (1994). The collective resource approach: The Scandinavian experience. *Scandinavian Journal of Information Systems, 6*(1), 71-84.

Kuhlthau, C.C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of American Society for Information Science, 42*(5), 361–371.

Kulturdepartementet. (1985). *Lov om folkebibliotek (folkebibliotekloven)* [The law on public libraries (the public library law]. Retrieved 1.3.2015 from https://lovdata.no/dokument/NL/lov/1985-12-20-108?q=folkebibliotek.

Kurniawan, S. & Conroy, G. V. (2007). Comparing comprehension speeds and accuracy of online information in students with and without dyslexia. In S. Kurniawan & P. Zaphiris (Eds.), *Advances in Universal Web Design and Evaluation* (pp. 257-270). Hershey: Idea Group.

Landerl, K., Fussenegger, B., Moll, K. & Willburger, E. (2009). Dyslexia and dyscalculia: Two learning disorders with different cognitive profiles. *Journal of Experimental Child Psychology, 103*(3), 309-324.

LanguageDaily. (n.d.). *1000 most common words in French*. Retrived 20.02.2013 from http://french.languagedaily.com/wordsandphrases/most-common-words.

LDA. (2008). *Learning disabilities vs. differences*. Retrieved 07.10.2015 from http://www.ldanys.org/index.php?s=2&b=25.

Lervåg, A., & Hulme, C. (2009). Rapid automatized naming (RAN) taps a mechanism that places constraints on the development of early reading fluency. *Psychological Science, 20*(8), 1040–1048.

Leuthold, S., Bargas-Avila, J. A. & Opwis, K. (2008). Beyond Web content accessibility guidelines: Design of enhanced text user interfaces for blind internet users. *International Journal of Human-Computer Studies, 66*(4), 257-270.

Lid, I. M. (2013). Developing the theoretical content in universal design. *Scandinavian Journal of Disability Research, 15*(3), 203-215.

Lindberg, T. & Näsänen, R. (2003). The effect of icon spacing and size on the speed of icon processing in the human visual system. *Displays, 24*(3), 111-120.

Logan, J. (2009). Dyslexic entrepreneurs: The incidence; their coping strategies and business skills. *Dyslexia, 15*(4), 328-346.

Lorigo et al. (2008). Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology, 59*(7), 1041-1052.

Lovdata. (2013). *Act relating to a prohibition against discrimination on the basis of disability (the Anti-Discrimination and Accessibility Act)*. Retrieved 17.09.2015 from http://app.uio.no/ub/ujur/oversatte-lover/data/lov-20130621-061-eng.pdf.

Lyster, S.-A. H. (2007). Reading development and reading disabilities: Focus on Norway. In T.E. Scruggs & M. Mastropieri (Eds.), *International Perspectives* (Advances in Learning and Behavioural Disabilities Vol. 20) (pp. 21–55). Bingley: Emerald Group Publishing.

Macdonald, S. J. (2009). Windows of reflection: Conceptualizing dyslexia using the social model of disability. *Dyslexia, 15*(4), 347-362.

MacFarlane, A., Al-Wabil, A., Marshall, C. R, Albrair, A., Jones, S. A. & Zaphiris, P. (2010). The effect of dyslexia on information retrieval: A pilot study. *Journal of Documentation, 66*(3), 307-326.

MacFarlane, A., Albrair, A., Marshall, C. R. & Buchanan, G. (2012). Phonological working memory impacts on information searching: An investigation of dyslexia. In J. Kamps, W. Kjraaj, N. Fuhr (Eds.) *Proceedings of IIiX 2012 the Fourth Information Interaction in Context Symposium, Nijmegen, the Netherlands, August 21-24, 2012* (pp 27-34). New York: ACM.

Maguire, M. (2001). Methods to support human-centred design. *International Journal of Human-Computer Studies, 55*(4), 587-634.

Malek, A., Hekmati, I., Amiri, S., Pirzadeh, J. & Gholizadeh, H. (2013). The standardization of Victoria Stroop Color-Word Test among Iranian Bilingual Adolescents. *Archives of Iranian Medicine, 16*(7), 380-384.

Mammarella, I. C., Pazzaglia, F. & Cornoldi, C. (2008). Evidence for different components in children's visuospatial working memory. *British Journal of Developmental Psychology, 26*(3), 337-355.

Mankoff, J., Fait, H., & Tran, T. (2005). Is your Web page accessible?: A comparative study of methods for assessing Web page accessibility for the Blind. In W. Kellogg, S. Zhai, G. v. d. Veer & C. Gale (Eds.), *CHI '05 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp 41-50). New York: ACM.

McCallum, M. L., & Bury, G. W. (2013). Google search patterns suggest declining interest in the environment. *Biodiversity and Conservation, 22*(6–7), 1355–1367.

McCarthy, J. E. & Swierenga, S. J. (2010). What we know about dyslexia and Web accessibility: A research review. *Universal Access in the Information Society, 9*(2), 147-152.

McGuire, J.M., Scott, S. S., & Shaw, S. F. (2006). Universal design and its applications in educational environments. *Remedial and Special Education, 27*(3), 166–175.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81-97.

Ministry of Children, Equality and Social Inclusion (2010). *Universell Utforming i Diskriminerings- og Tilgjengelighetsloven* [Universal Design in the Anti-Discrimination and Accessibility Act]. Retrieved 4.12.2015 from https://www.regjeringen.no/globalassets/upload/BLD/Universellutforming/RundskrivDiskriminering.pdf

Moats, L.C. (1996). Phonological spelling errors in the writing of dyslexic adolescents. *Reading and Writing, 8*(1), 105–119.

Moores, E., Cassim, R. & Talcott, J. B. (2011). Adults with dyslexia exhibit large effects of crowding, increased dependence on cues, and detrimental effects of distractors in visual search tasks. *Neuropsychologica, 49*(14), 3881-3890.

Moreno, L., Martínez, P. & Ruiz-Mezcua, B. (2008). Disability standards for multimedia on the Web. *MultiMedia, IEEE, 15*(4), 52-54.

Mortimore, T. & Crozier, W. R. (2006). Dyslexia and difficulties with study skills in higher education. *Studies in Higher Education,* 31(2), 235-251.

Moyer, J. E. (2012). Audiobooks and e-books: A literature review. *Reference & User Service Quarterly, 51*(4), 340-354.

Mueller, S. T. (2010). *The Psychology Experiment Building Language, Version 0.11*. Retrieved 10.08.2012 from http://pebl.sourceforge.net.

Mueller, S. T. & Piper, B. J. (2014). The Psychology Experiment Building Language (PEBL) and PEBL Test Battery. *Journal of Neuroscience Methods, 222*, 250-259.

Neisser, U. (1964). Visual search. *Scientific American, 210*(6), 94-102.

Neuhaus, G., Foorman, B. R., Francis, D. J. & Carlson, C. D. (2001). Measures of information processing in rapid automatized naming (RAN) and their relation to reading. *Journal of Experimental Child Pscyhology, 78*(4), 359-373.

Newell, A. F. (1995). Extra-ordinary human-computer interaction. In A. Edwards (Ed.), *Extra-ordinary human-computer interaction: Interfaces for users with disabilities* (pp. 3-18). Cambridge: Cambridge University Press.

Newell, A. F. & Gregor, P. (2000). "User Sensitive Inclusive Design": In search of a new paradigm. In J. Thomas (Ed.), *CUU '00 Proceedings on the 2000 conference on Universal Usability* (pp. 39-44). New York: ACM.

Nicolson, R. I. & Fawcett, A. J. (1990). Automaticity: A new framework for dyslexia research?. *Cognition, 35*(2), 159-182.

Nicolson, R. I. & Fawcett, A. J. (2011). Dyslexia, dysgraphia, procedural learning and the cerebellum. *Cortex, 47*(1), 117-127.

Nielsen, J. (1990). Miniatures versus icons as visual cache for videotext browsing. *Behaviour & Information Technology, 9*(6), 441-449.

Nielsen, J. (1994). Heuristic evaluation. In J. Nielsen & R. L. Mack (Eds.), *Usability inspection methods* (pp.25-62). New York: Wiley.

Nielsen, J. & Levy, J. (1994). Measuring usability: Preference vs. Performance. *Communications of the ACM, 37*(4), 66-75.

Niemelä, N. & Saarinen, J. (2000). Visual search for grouped versus ungrouped icons in a computer interface. *Human Factors, 42*(4), 630-635.

Norman, D. (2002). *The design of everyday things*. New York: Basic Books.

Norwegian Archive, Library and Museum Authority. (2004). *Er biblioteket tilgjengelig for alle?* [Is the library accessible for all?]. Retrieved 17.07.2016 from http://www.universell-utforming.miljo.no/file_upload/det_tilgjengelige_bibl_sjekk.pdf

Nothdurft, H.-C. (1999). Focal attention in visual search. *Vision Research, 39*(14), 2305-2310.

NSD. (2016). *Database for statistikk om høgere utdanning: Registrerte studenter* [Database on statistics about higher education: Registred students]. Retrieved 1.3.2015 from http://dbh.nsd.uib.no/statistikk/rapport.action?visningId=124&visKode=false&columns=arstall!8!kjonn&index=1&formel=222&hier=insttype!9!instkode!9!fakkode!9!ufakkode!9!progkode&sti=&param=arstall%3D2015!9!dep_id%3D1!9!kategori%3DS!9!semester%3D3

Ojanpää. H., Näsänen, R. & Kojo, I. (2002). Eye movements in the visual search of word lists. *Vision Research, 42*(12), 1499-1512.

Oliver, M. (1996). *Understanding disability: From theory to practise*. New York: Palgrave.

Oxford Dictionaries. (n.d.). *The OEC: Facts about the language*. Retrieved 20.02.2013 from http://oxforddictionaries.com/words/the-oec-facts-about-the-language.

Pagulayan, K. F., Busch, R. M., Medina, K. L., Bartok, J. A. & Krikorian, R. (2006). Developmental normative data for the Corsi Block-Tapping Task. *Journal of Clinical & Experimental Neuropsychology, 28*(6), 1043-1052.

Palmer, S. (2000). Phonological recording deficit in working memory of dyslexic teenagers. *Journal of Research in Reading, 23*(1), 28-40.

Pan, B., Hembrooke, H. A., Gay, G. K., Granka, L. A., Feusner, M. K. & Newman, J. K. (2004). The determinants of Web page viewing behaviour: An eye-tracking study. In A. T. Duchowski, R. Vertegaal (Eds.), *ETRA '04 Proceedings of the 2004 Symposium on Eye tracking research & applications* (pp. 147-154). New York: ACM.

Pelli, D. G. & Tillman, K. E. (2008). The uncrowded window of object recognition. *Nature Neuroscience, 11*(10), 1129-1135.

Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E. & McCarley, J. S. (2001). Visual search has memory. *Psychological Science, 12*(4), 287-292.

Petrie, H. & Bevan, N. (2009). The evaluation of accessibility, usability and user experience. In C. Stephanidis (Ed.), *The Universal Access Handbook (Human Factors and Ergonomics* (pp. 147-154). Boca Raton: CRC Press.

Petrie, H., Hamilton, F., King, N. & Pavan, P. (2006). Remote usability evaluations with disabled people. In R. Grinter, T. Rodden, P. Aoki, E. Cutrell, R. Jeffries, G. Olson (Eds.), *CHI ’06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1133-1141). New York: ACM.

Petrie, H. & Kheir, O. (2007). The relationship between accessibility and usability of Websites. In M. B. Rosson & D. Gilmore (Eds.), *CHI ‘07 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 397-406). New York: ACM.

Phillips, M. H. & Edelman, J. A. (2008). The dependence of visual scanning performance on saccade, fixation, and perceptual metrics. *Vision Research, 48*(7), 926-936.

Poole, A. & Ball, L. J. (2006). Eye tracking in HCI and usability research. In: C. Ghaoui (Ed.), *Encyclopedia of Human Computer Interaction* (pp. 211-219). Hershey: Idea Group.

Prado, C., Dubois, M. & Valdois, S. (2007). The eye movements of dyslexic children during reading and visual search: Impact of the visual attention span. *Vision Research, 47*(19), 2521-2530.

Purcell, K., Brenner, J. & Rainie, L. (2012). *Search engine use 2012*. Retrieved 18.09.2015 from http://www.pewinternet.org/files/old-media/Files/Reports/2012/PIP_Search_Engine_Use_2012.pdf.

Rabbetts, R. B. (2007). *Clinical visual optics*. Edinburgh: Elsevier.

Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or sensorimotor dysfunction?. *Current Opinion in Neurobiology, 13*(2), 212-218.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457-1506.

Reich, L. N. & Ekabutr, M. (2012). The effects of optical defocus on the legibility of the Tumbling-E and Landolt-C. *Optometry and vision science, 79*(6), 389-393.

Rele, R. S. & Duchowski, A. T. (2005). Using eye tracking to evaluate alternative search results interfaces. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 49(15), 1459-1463.

Rello, L., & Baeza-Yates, R. (2013). Good fonts for dyslexia. In *ASSESTS ’13 Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (article no. 14). New York: ACM.

Rello, L., Kanvinde, G. & Baeza-Yates, R. (2012). Layout guidelines for web text and a web service to improve accessibility for dyslexics. In M. Vigo, J. Abascal, R. Lopes & P. Salomoni (Eds.), *W4A’12 Proceedings of the International Cross-Disciplinary Conference on Web Accessibility* (article no. 36). New York: ACM.

Rello, L., Pielot, M., Marcos, M.-C., & Carlini, R. (2013). Size matters (spacing not): 18 points for a dyslexic-friendly Wikipedia. In G. Brajnik & P. Salomoni (Eds.), *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility, May 13–15, 2013, Rio de Janeiro, Brazil (W4A’13)* (article no. 17). New York: ACM.

Ricci, F., Cedrone, C. & Cerulli, L. (1998). Standardized measurement of visual acuity. *Opthalmic Epidemology, 5*(1), 41-53.

Richards, J. T. & Hanson, V. L. (2004). Web accessibility: A broader view. In S. Feldman, M. Uretsky, M. Najork & C. Wills (Eds.), *WWW ’04: Proceedings of the 13th International Conference on World Wide Web* (pp 72-79). ACM New York

Riddick, B. (1996). *Living with dyslexia: The social and emotional consequences of specific learning difficulties*. London: Routledge.

Riddick, B., Farmer, M. & Sterling, C. (1997). *Students and dyslexia: Growing up with a specific learning difficulty*. London: Whurr.

Riddick, B., Sterling, C., Farmer, M. & Morgan, S. (1999). Self-esteem and anxiety in the educational histories of adult dyslexic students. *Dyslexia, 5*(4), 227-248.

Rupprecht, D., Blum, R. & Bomsdorf, B. (2014). User centered inclusive design for people with dyslexia: Experiences from a project on accessibility. In S. Sauer, C. Bogdan, P. Forbig, R. Bernhaupt & M. Winckler (Eds.), *Human-Centered Software Engineering, 5th IFIP WG 13.2. International Conference, HCSE 2014, Parderborn, Germany, September 16-18, 2014* (pp. 307-314). Heidelberg: Springer.

Rutledge, H. (2002). Dyslexia: Challenges and opportunities for public libraries. *Journal of Librarianship and Information Science, 34*(3), 134-144.

Rømen, D. & Svanæs, D. (2012). Validating WCAG versions 1.0 and 2.0 through usability testing with disabled users. *Universal Access in the Information Society, 11*(4), 375-385.

Sahib, N. G., Tombros, A. & Stockman, T. (2012). A comparative analysis of the information-seeking behavior of visually impaired and sighted searchers. *Journal of the American Society for Information Science, 63*(2), 377-391.

Samuelsson, S., Herkner, B. & Lundberg, I. (2003). Reading and writing difficulties among prison inmates: A matter of experimental factors rather than dyslexic problems. *Scientific Studies of Reading, 7*(1), 53-73.

Santana, V. F. de, Oliveira, R. de., Almeida, L. D. A., & Baranauskas, M. C. C. (2012). Web accessibility and people with dyslexia: A survey on techniques and guidelines. In M. Vigo, J. Abascal, R. Lopes, & P. Salomoni (Eds.), *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility (W4A'12) Lyon, France, 16–17 April* (article no.35). New York: ACM.

Schmetzke, A. (2006). Online distance education: "Anytime, anywhere" but not for everyone. *Information Technology and Disability Journal, 8*(2).

Schneps, M. H., Rose, L. T. & Fischer, K. W. (2007). Visual learning and the brain: Implications for dyslexia. *Mind, Brain and Education, 1*(3), 128-139.

Schuchardt, K., Maehler, C. & Hasselhorn, M. (2008). Working memory deficits in children with specific learning disorders. *Journal of Learning Disabilities, 41*(6), 514-523.

Schwartz, C. (1998). Web Search Engines. *Journal of the American Society for Information Science, 49*(11), 973-982.

Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). *Experimental and quasi-experimental design for generalized causal inference*. Belmont: Wadsworth.

Shakespeare, T. (2004). Social models of disability and other life strategies. *Scandinavian Journal of Disability Research, 6*(1), 8-21.

Shakespeare, T. (2013). The social model of disability. In L. J. Davis (Ed.), *The Disability Studies Reader* (pp. 214-221). New York: Routledge.

Shaywitz, S. E. & Shaywitz, B. A. (2001). The neurobiology of dyslexia. *Focus on Basics, 2001*, 11-15.

Shaywitz, S. E. & Shaywitz, B. A. (2005). Dyslexia (specific reading disability). *Biological Psychiatry, 57*(11), 1301–1309.

Siegel, L. S. & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60*(4), 973-980.

Sireteanu, R., Goebel, C., Goertz, R., Werner, I., Nalewajko, M. & Thiel, A. (2008). Impaired serial visual search in children with developmental dyslexia. *Annals of the New York Academy of Sciences, 1145*, 199-211.

Skottun, B. C. & Skoyles, J. (2007). The use of visual search to assess attention. *Clinical and Experimental Optometry, 90*(1), 20-25.

Smith-Spark, J. H. & Fisk, J. E. (2007). Working memory functioning in developmental dyslexia. *Memory, 15*(1), 34-56.

Smythe, I. (2011). Dyslexia. *British Journal of Hospital Medicine, 72*(1), 39-43.

Snowling, M. J. (2000). *Dyslexia*. Malden: Blackwell.

Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia, 7*(1), 37-46.

Snowling, M. J., Duff, F., Petrou, A., Schieffledrin, J. & Bailey, A. M. (2011). Identification of children at risk of dyslexia: The validity of teacher judgements using 'Phonic Phases'. *Journal of Research in Reading, 34*(2), 157-170.

Solheim, O. J. & Uppstad, P. H. (2011). Eye-tracking as a tool in process-oriented reading test validation. *International Electronic Journal of Elementary Education, 4*(1), 153-168.

Sosialdepartement. (2003). *Nedbygging av funksjonshemmende barrierer* [Dismantling disabling barriers]. Retrieved 16.12.2015 from https://www.regjeringen.no/contentassets/5a8122df4dee44a38beb1ca42698b490/no/pdfs/stm200220030040000dddpdfs.pdf

Spinuzzi, C. (2005). The methodology of participatory design. *Technical Communication, 52*(2), 163-174.

Stein, J. (2001). The magnocellular theory of developmental dyslexia. *Dyslexia, 7*, 12-36.

Story, M. F. (1998). Maximizing usability: The principles of universal design. *Assistive Technology, 10*(1), 4-12.

Story, M. F., Mueller, J. L. & Mace, R. L. (1998). *The universal design file: Designing for people of all ages & abilities*. Raleigh: Center for Universal Design.

Strauss, E., Sherman, E. M. S. & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. Oxford: Oxford University Press.

Sutton, L. B., Erlen, J. A., Glad, J. A. M. & Siminoff, L. A. (2003). Recruiting vulnerable population for research: Revisiting the ethical issues. *Journal of Professional Nursing, 19*(2), 106-112.

Taylor, R. S. (1968). Question-Negotiation and information seeking in libraries. *College & Research Libraries, 29*(3), 178-194.

The accessible library. (n.d.). *Accessible ICT*. Retrieved 17.07.2016 from http://www.of.fylkesbibl.no/ostfyb/tb/elektroniskinfo.html.

Tops, W., Callens, M., Lammertyn, J., Hees, V. V. & Brysbaert, M. (2012). Identifying students with dyslexia in higher education. *Annals of Dyslexia, 62*(3), 186-203.

Treasury Board of Canada. (2013). *Standard on Web Accessibility*. Retrieved 6.1.2016 from http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?section=text&id=23601

Treisman, A. M. & Gelade, G. (1980). A Feature-Integration Theory of Attention. *Cognitive Psychology, 12*(1), 97-136.

Turetken, O. & Sharda, R. (2005). Clustering-based visual interfaces for presentation of Web search results: An empirical investigation. *Information Systems Frontiers, 7*(3), 273-297.

UN. (2006). *Convention on the rights of persons with disabilities*. Retrieved 17.09.2015 from http://www.un.org/disabilities/convention/conventionfull.shtml.

Universell. (2008). *Brukbarhetspyramiden* [The usability pyramid]. Retrieved 1.12.2015 from http://uukurs.universell.no/gammelt-innhold/universell-utforming/brukbarhetspyramiden/

Van der Bosch, A., Bogers, T. & de Kunder, M. (2016). Estimating search engine index size variability: A 9-year longitudinal study. *Scientometrics, February*, 1-18.

Van der Schoot, M., Licht, R., Horsley, T. M. & Sergeant, J. A. (2000). Inhibitory deficits in reading disability depend on subtype: Guessers not spellers. *Child Neuropsychology, 6*(4), 297-312.

Vandierendonck, A., Kemps, E., Fastame, M. C. & Szmalec, A. (2004). Working memory components of the Corsi Block tasks. *British Journal of Psychology, 95*(1), 57-79.

Vidyasagar, T. R. & Pammer, K. (1999). Impaired visual search in dyslexia relates to the role of the magnocellular patway of attention. *NeuroReport, 10*(6), 1283-1287.

Vitu, F., McConkie, G. W., Kerr, P. & O'Reagan, J. K. (2001). Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision Research, 41*(25), 3513-3533.

Vlaskamp, B. N. S. & Hooge, I. Th. C. (2006). Crowding decreases saccadic search performance. *Vision Research, 46*(3), 417-425.

W3C. (1999). *Web Content Accessibility Guidelines 1.0*. Retrieved 28.09.2015 from http://www.w3.org/TR/WAI-WEBCONTENT/.

W3C. (2005). *Introduction to Web accessibility*. Retrieved 15.12.2015 from https://www.w3.org/WAI/intro/accessibility.php

W3C. (2008). *Web content accessibility guidelines (WCAG) 2.0*. Retrieved 16.09.2015 from http://www.w3.org/TR/WCAG20/.

W3C. (2014). *Web Accessibility Initiative*. Retrieved 17.11.2015 from http://www.w3.org/WAI/.

W3C. (2015). *Contrast (Minimum): Understanding SC 1.4.3*. Retrieved 17.12.2015 from https://www.w3.org/TR/UNDERSTANDING-WCAG20/visual-audio-contrast-contrast.html

Ward, D., Hahn, J. & Feist, K. (2012). Autocomplete as research tool: A study on providing search suggestions. *Information Technology and Libraries, 31*(4), 6-19.

Ward, R. (2001). Visual attention moves no faster than the eyes. In: K. Shapiro (Ed.), *The limits of attention: Temporal Constraints in Human Information Processing* (pp 199-216). Oxford: Oxford University Press.

Webb, N. & Renshaw, T. (2008). Cognitive modelling in HCI research. In P. Cairns & A. L. Cox (Eds.), *Research Methods for Human-Computer Interaction*. Cambridge: Cambridge.

Weisen, M., Petrie, H., King, N. & Hailton, F. (2005). Web accessibility revealed: The Museums, Libraries and Archives Council audit. *Ariadne*, *44*.

White, R. W. & Marchionini, G. (2007). Examining the effectiveness of real-time query expansion. *Information Processing and Management, 43*(3), 685-704.

White, R. W. & Morris, D. (2007). Investigating the querying and browsing behavior of advanced search engine users. In W. Kraaj & A. P. De Vries (Eds.), *SIGIR '07 Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 255-262). New York: ACM.

Williams, & Hennig (2015a). Effect of Web page menu orientation on retrieving information by people with learning disabilities. *Journal of the Association for Information Science and Technology, 66*(4), 674-683.

Williams, P. & Hennig, C. (2015b). Optimising Web site designs for people with learning disabilities. *Journal of Research in Special Education Needs, 15*(1), 25-36.

Williamson, K. & Schauder, K. (2000). Information seeking by blind and sight impaired citizens: An ecological study. *Information Research, 5*(4), no pagination.

Wilson, T. D. (1999). Models in information behaviour research. *Journal of Documentation, 55*(3), 249-270.

Wilson, T. D. (2000). Human information behavior. *Informing Science, 3*(2), 49-55.

Wolfe, J., Cave, K. & Franzel, S. (1989). Guided search: An alternative to the feature integrated model for visual search. *Journal of Experimental Psychology: Human, Perception and Performance, 15*(3), 419-433.

Wolff, U. & Lundberg, I. (2002). The prevalence of dyslexia among art students. *Dyslexia, 8*, 34-42.

Wolff, U. & Lundberg, I. (2003). A technique for group screening of dyslexia among adults. *Annals of Dyslexia, 53*, 324-339.

Wright, C. M., Conlon, E. G. & Dyck, M. (2012). Visual search deficits are independent of magnocellular deficits in dyslexia. *Annals of Dyslexia, 62*(1), 53-69.

Wright, R. D. & Ward, L. M. (2008). *Orienting of attention*. Oxford: Oxford University Press.

Zaphiris, P., Kurniawan, S. & Ghiawadwala, M. (2007). A systematic approach to the development of research-based Web design guidelines for older people. *Universal Access in the Information Society, 6*(1), 59-75.

Zarach, V. (2002). *Ten guidelines for improving accessibility for people with dyslexia*. Retrieved 14.2.2016 from http://wiki.cetis.ac.uk/Ten_Guidelines_for_Improving_Accessibility_for_People_with_Dyslexia

Zhang, H. & Salvendy, G. (2001). The implications of visualization ability and structure preview design for Web information search tasks. *International Journal of Human-Computer Interaction, 13*(1), 75-95.

Zhang, P., Bobier, W., Thompson, B. & Hess, R. F. (2011). Binocular balance in normal vision and its modulation by mean luminance. *Optometry & Vision Science, 88*(9), 1072-1079.

Ziegler, J. C., Castel, C., Pech-Georgel, C., George, F., Alario, F-X. & Perry, C. (2008). Developmental dyslexia and the dual route model of reading: Simulating individual differences and subtypes. *Cognition, 107*(1), 151-178.

Ziegler, J. C. & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*(1), 3-29.

Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M. … Ziegler, J. C. (2012). Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences of the United States of America, 109*(28), 11455-11459.

Appendix A

Title:	The effect of dyslexia on searching visual and textual content: Are icons really useful?
Authors:	Gerd Berget and Frode Eika Sandnes
Proceeding:	M. Antona & C. Stephanidis (Eds.), Universal Access in Human-Computer Interaction: Access to Learning, Health and Well-being: 9th International Conference, UAHCI 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings, Part III
Publisher:	Springer
Year:	2015
Pages:	616-625

Appendix B

Title: Is visual content in textual search interfaces beneficial to dyslexic users?

Authors: Gerd Berget, Fiona Mulvey and Frode Eika Sandnes

Journal: International Journal of Human-Computer Studies

Volume: 92-93

Year: 2016

Pages: 17-29

Appendix C

Title: Searching databases without query-building aids: Implications for dyslexic users

Authors: Gerd Berget and Frode Eika Sandnes

Journal: Information Research

Volume: 20

Issue: 4

Year: 2015

Pages: paper 689

informationresearch
VOL. 20 NO. 4, DECEMBER, 2015

Contents | Author index | Subject index | Search | Home

Searching databases without query-building aids: implications for dyslexic users

Gerd Berget
Institute of Information Technology, Faculty of Technology, Art and Design, Oslo and Akershus University College of Applied Sciences, Oslo, Norway
Frode Eika Sandnes
Faculty of Technology, Westerdals Oslo School of Arts, Communication and Technology. Oslo, Norway

Abstract

Introduction. *Few studies document the information searching behaviour of users with cognitive impairments. This paper therefore addresses the effect of dyslexia on information searching in a database with no tolerance for spelling errors and no query-building aids. The purpose was to identify effective search interface design guidelines that benefit dyslexic users.*
Method. *Twenty dyslexic students and twenty controls solved ten predefined search tasks in the Norwegian library catalogue Bibsys Ask. Screen-recording and eye-tracking were used to observe search behaviour.*
Analysis. *The statistical analysis software SPSS was used for quantitative analyses, SMI BeGaze was used for qualitative analysis of the search behaviour.*
Results. *Dyslexic students took longer and formulated more queries per task than the controls. Further, they submitted shorter queries, made more misspellings and relied more upon external resources. There were no differences in problem solving approaches across the two groups except that the dyslexic students used external Websites more.*
Conclusions. *The results of this study indicate that dyslexia has a negative impact on search performance in systems with no tolerance for errors and no query-building aids. Several guidelines are suggested based on the observed information searching behaviour to accommodate users with dyslexia.*

change font

Introduction

Online databases are important sources for scholarly information. There is a growing reliance among students on search engines such as Google (Jamali and Asadi, 2010). General search engines often have basic search interfaces as the default option, where the user inputs all the search terms in one search field. Further, these systems often offer full text search, allowing the users to search with natural language. Search engines may also have a high tolerance for spelling errors and offer various query-building aids. Autocomplete is an example of a query-building aid where query-terms are suggested while the user is entering the query. Query-building aids may also return relevant results even if queries are misspelled.

General search engines do not fully accommodate the needs of students and scholars. Additional resources such as library catalogues and article databases are often required to locate articles on reading lists and to find relevant citations for written assignments. Domain specific search systems often contain advanced interfaces with several search fields, which require different content such as author names, titles or keywords. Moreover, these databases do often not offer full text search. Although a database may provide synonym control, there are higher demands on users to input query terms matching the vocabulary of the database. Bibliographic databases often have a low or no tolerance for spelling errors and offer no query-building aids. Such systems require well-developed writing and reading skills, which are typically impaired in dyslexic users. It is suggested that students with impairments such as dyslexia struggle with information seeking (Hepworth, 2007; "MacFarlane, Al-Wabil, Marshall, Albrair, Jones and Zaphiris, 2010; MacFarlane, Albrair, Marshall, and Buchanan, 2012; Habib *et al.*, 2012). Dyslexics may also have difficulties in using physical libraries, and thus rely on extra library support (Mortimore and Crozier, 2006).

The number of dyslexic students entering higher education is increasing (Tops, Callens, Lammertyn, Van Hees and Brysbaert, 2012). Information literacy is often emphasised as a strategy for lifelong learning (Association of College and Research Libraries, 2000) and access to scholarly literature is important to this end. Inaccessible search systems may be an obstacle in completing a higher education qualification for dyslexic students. Consequently, there is a need for more accessible search systems, but to date little is known about the information searching behaviour of dyslexic users in order to facilitate these more accessible search systems (Hepworth, 2007; MacFarlane *et al.*, 2012). The novel contribution of this study is to provide new empirical evidence on how dyslexia affects information search. For example, do dyslexic students take longer and formulate more queries per task than the students without dyslexia, do dyslexics submit shorter queries, make more misspellings and rely more on external resources than users without dyslexia. Such empirical knowledge is needed to develop design guidelines which better accommodate search experiences for dyslexic users.

This study involved forty students (twenty with dyslexia and twenty controls) who each solved ten predefined search tasks in the Web-based Norwegian library catalogue Bibsys Ask. The null hypothesis herein was that dyslexia would not affect overall search performance. It was expected that several search parameters would be affected by dyslexia, which typically causes spelling and reading difficulties (Snowling, 2000). It was assumed that spelling skills would affect the time taken to pursue a query, the number of queries needed to solve a task, query lengths and portion of misspellings. Further, reading skills could affect query lengths, since it was assumed that slow readers would formulate queries with a higher precision level to reduce the effort needed to browse result lists. It was also assumed that a broad range of problem-solving approaches would be applied. The following hypotheses were the basis for the study:

> ***H1:*** *Dyslexic users take more time formulating queries than controls.*
> ***H2:*** *Dyslexic users formulate more queries per task than controls.*
> ***H3:*** *Dyslexic users submit shorter queries than controls.*
> ***H4:*** *Dyslexic users produce more misspellings than controls.*
> ***H5:*** *Dyslexic users apply a broader range of problem solving strategies than controls.*

Background

Dyslexia is a learning disorder which affects 3-10% of the general population (Snowling, 2000). Although the etiology is disputed (Jones, Branigan and Kelly, 2007; Schuchardt, Maehler and Hasselhorn, 2008), dyslexia is commonly connected to an underlying phonological deficit (Snowling, 2001). Dyslexia typically affects reading and writing activities however impairments in word retrieval, working memory, concentration and rapid naming skills are also prevailing (Snowling, 2000; Snowling, 2001; Mortimore and Crozier, 2006; Smith-Spark and Fisk, 2007). Rapid naming skills concern the ability to connect visual and oral information by assigning appropriate names to, for instance, numbers, colours or objects (Lervåg and Hulme, 2009). Difficulties related to dyslexia persist until adulthood and remain throughout life (Swanson and Hsieh, 2009).

Information searching involves a variety of cognitive skills. For instance, rapid naming skills and word retrieval are used to formulate effective search queries. Spelling skills are needed to input terms correctly (MacFarlane, Albrair, Marshall, and Buchanan, 2012). Reading skills and a good working memory are required when browsing result lists and assessing documents. Several of these skills are often impaired in dyslexic users. Consequently, dyslexia may have a negative effect on search performance, especially in systems with no query-building aids and no tolerance for spelling errors.

The field of human computer interaction has become increasingly concerned with user diversity and the importance of accessible interfaces (Henry, Abou-Zahra and Brewer, 2014). Universal design is included in several national legislations (Imrie, 2012). The Web Content Accessibility Guidelines (WCAG) is often referred to in this context. However, these guidelines may not meet the needs of users with various disabilities (Rømen and Svanæs, 2012) or users in cognitive language and learning areas (W3C, 2008) such as dyslexia (Freire, Petrie and Power, 2011; deSantana, Oliveira, Almeida and Baranauskas, 2012). Further, existing guidelines typically address general Website design, and not specific issues related to search systems.

More knowledge about user behaviour is needed to develop effective interface guidelines. Dyslexics have received little attention in studies on Web accessibility and search systems (de Santana, Oliveira, Almeida and Baranauskas, 2012; McCarthy & Swierenga, 2010). Furthermore, most research on dyslexia and Website design address navigation (Al-Wabil, Zaphiris and Wilson, 2007), font types, font sizes (Rello and Baeza-Yates, 2013; Rello, Pielot, Marcos and Carlini, 2013) and accessible textual content (Evett and Brown, 2005). Consequently, few documented studies exist about information retrieval and the actual use of search interfaces by dyslexic users.

General user-centred theories and models on information seeking behaviour do generally not address users with impairments. Some studies address blind library users, for instance Craven & Brophy (2003) and Sahib, Tombros and Stockman (2012), while cognitive disabilities have received less attention (MacFarlane, Al-Wabil, Marshall, Albrair, Jones and Zaphiris, 2010). Of the few studies that addressed general cognitive abilities as possible predictors of information searching Kim and Allen (2002) investigated whether cognitive abilities such as different levels of perceptual speed, spatial scanning and logical reasoning affected the information seeking behaviour of students searching the Web. Other studies discuss how cognitive style or learning style affects the retrieval process, for instance Palmquist and Kim (2000), Papaeconomou, Zijlema and Ingwersen (2008) and Al-Maskari and Sanderson (2011).

Of the few studies that addressed how dyslexia affects information searching behaviour MacFarlane et al. (2010) compared the search behaviour of 5 dyslexics and 5 controls in the test retrieval system Okapi. The results from this study were that dyslexic users on average exhibited fewer search iterations, took more time on each search and reviewed fewer documents than the controls. The authors found no misspellings and could therefore not explain how such mistakes affect query variables. A later study by MacFarlane et al. (2012) investigated the search behaviour of 8 dyslexics and 8 controls with a focus on the phonological working memory. The phonological working memory includes spoken and written material and is often impaired in users with dyslexia (Smith-Spark and Fisk, 2007). Participants in the control group classified more documents as irrelevant than the dyslexia group. Furthermore, the authors found a correlation between working memory and the ratio of documents classified as irrelevant. The conclusion was that there is a need for more knowledge on the information searching behaviour of users with dyslexia.

Berget and Sandnes (2015) observed the information searching behaviour of 20 dyslexics and 20 controls solving ten predefined search tasks in the general search engine Google with query building aids. Although the dyslexics made more misspellings, there were no significant differences in time usage or number of queries per task compared to controls. The dyslexic students fixated significantly less on the screen and the autocomplete suggestions while entering queries than the controls. Berget and Sandnes (2015) concluded that a high tolerance for spelling errors may be as useful as autocomplete functions for dyslexic users, because of the extensive focus on the keyboard during text entry. Results from this study also indicate that help functions such as displaying a red line beneath misspelled words in the search field, may be helpful to users.

Method

Participants

The study included 40 volunteering students, 20 diagnosed with dyslexia and 20 controls. The participants were recruited through advertisements by the interest association Dyslexia Norway, announcements in classrooms and on Websites. Four of the participants (three dyslexics and one control) were diagnosed with Attention Deficit Hyperactivity Disorder (ADHD) or Attention Deficit Disorder (ADD). ADHD and ADD are typically characterized by short attention span, excessive activity and impulsivity, and comorbidity with dyslexia is well documented (Germanò, Gagliano, and Curatolo, 2010). Participants in the control group were matched against the dyslexia group according to gender, age, field of study and year of study. Prevalence of dyslexia was the only controlled variable.

The participants included 60% females and 40% males, aged 19-40 years with a mean age of 24.2 years (*SD*=5.2) in the dyslexia group and 23.6 (*SD*=3.8) in the control group. Dyslexia is typically more prevalent among men than women (Hawke, Olson, Willcut, Wadsworth and DeFries, 2009). Consequently, the portion of male participants should have been higher. However, the skewed gender distribution was caused by recruiting difficulties.

The students were enrolled onto either three-year bachelor-programmes or two-year master-programmes, with mean year of study of 2.4, for both the dyslexia group (*SD*=1.2) and the control group (*SD*=1.0) (value 1-3 used for bachelor students, 4-5 for master students). The participants represented a diversity of disciplines and study programmes, for instance nursing, engineering, social sciences and humanities. Students from library- and information science programmes were excluded because of their extensive training in information searching. None of the dyslexic students regularly relied on assistive technologies. All participants were regarded as high-functioning dyslexics since they successfully have been admitted to higher degree programmes and have demonstrated academic progress. This may be a limitation of the study, since other groups with dyslexia may have less academic training.

Formal training and prior experience with search systems may affect information searching (Kuhlthau, 1991). Eight dyslexic participants had received formal training in searching library catalogues compared to nine users in the control group. The controls had slightly more years of experience than the dyslexics, both in using general search engines (M=14.4, SD=3.2 compared to M=13.9, SD=3.3) and library catalogues (M=4.6, SD=3.9 compared to M=3.9, SD=3.1). The students with dyslexia reported a higher frequency of use of the library catalogues than the control group.

All participants were screened for dyslexia using a word chain test (Høien and Tønnesen, 2008). The dyslexic students had previously been diagnosed by psychological or pedagogical professionals. However, the screening test eliminated the need for consulting sensitive medical records to confirm the dyslexia diagnosis and ensured that no controls were dyslexic. The word chain test is a standardised test in Norway which is applied in previous research, for instance Solheim and Uppstad (2011) and Dahle and Knivsberg (2014). A low score is considered indicative of dyslexia and diagnostic tests are recommended for adult scores below 43 (Høien and Tønnesen, 2008). The dyslexic students scored significantly lower (M=39.3, SD=10.3) than the controls (M=60.2, SD=10.0), $t(38)=6.52$, $p<.001$, $d=2.12$.

Visual acuity was measured, since reduced vision can be a source of error. Landolt C charts were used at 4 m for distance and 40 cm for near, in accordance with the European standard (ISO, 2009). All participants had at least acuity of 0.6 on each eye separately and 0.8 with both eyes open for both tests, which is considered within the limits of normal visual acuity (Zhang, Bobier, Thompson and Hess, 2011).

Procedure

The participants were instructed to solve ten predefined search tasks using the library catalogue. Instructions and search tasks were presented orally using pre-recorded speech synthesis files. Oral instructions may prevent misinterpretations of tasks, since dyslexia may affect reading speed and reading comprehension (Ransby and Swanson, 2003). Moreover, oral instructions prevented the participants from seeing the spelling of terms. The experimental design choice of introducing oral instructions places more cognitive load on the participants since they have to remember the tasks. This could be a challenge since dyslexia is also often associated with reduced working memory. However, the task of reading the instructions was considered more cognitively demanding than memorising the short oral instructions. The assumed discrepancy between reading and listening abilities among individuals with reading disabilities is also supported by the literature (Badian, 1999).

An image representing the question was displayed while instructions were given (Figure 1). The task number allowed the participants to track the progress of the session.

Figure 1: Example stimulus, task ten

At the start of the session the participants were asked about prior experiences with online library catalogues including Bibsys Ask. All ten tasks were completed in one session, and there was no dialogue between the participants and the experimenter besides responses to questions initiated by the participants. The students were allowed to use external online resources as supplementary aids if needed e.g. checking the spelling of a word. They were asked to comment on their search behaviour during debriefing where unusual behaviour had been observed.

Ethics

The study was ethically screened and approved by the Norwegian Social Science Data Services (project number 29348) and all data was anonymised. Moreover, all participants signed consent forms. None of the students were related to the researchers.

Apparatus

The experiments were run in the visual stimulus presentation software SMI Experiment Center version 3.2.11 using the background screen recorder. The stimulus was displayed on a 21' Dell LED flat screen with the resolution set to 1680 x 1050 pixels. Windows Internet Explorer version 9 was used for the search tasks, with all functions related to prior use such as browsing history and cookies disabled. All searches were conducted in the basic search mode in Bibsys Ask (Figure 2). Data were analysed using SMI BeGaze version 3.2 and the statistical analysis software IBM SPSS Statistics 22.

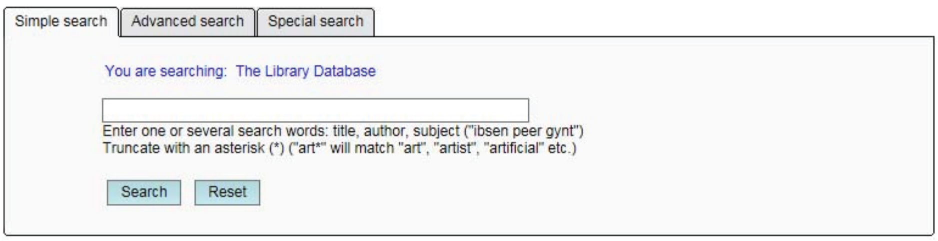


Figure 2: Simple search mode in Bibsys Ask (http://ask.bibsys.no/ask/action/smpsearch), reproduced with permission. The Norwegian interface was used in the study.

Materials

Bibsys Ask is a bibliographic library database used by all the higher education libraries in Norway, in addition to several research libraries and institutions. This database has no tolerance for spelling errors and provides no query-building aids; the system merely instructs the user to check the spelling or modify the query if no results are returned (Figure 3).

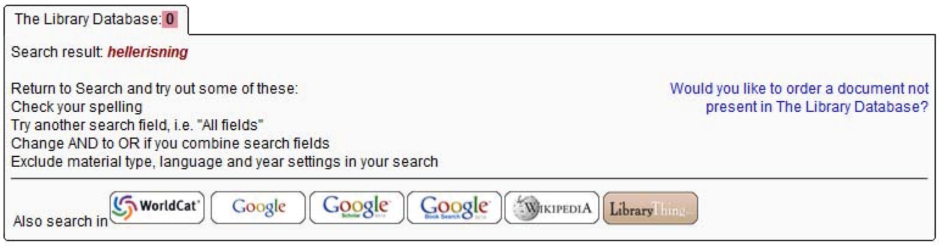


Figure 3: Instructions from Bibsys Ask when no results are returned, reproduced with permission

The participants were given ten predefined search tasks in Norwegian (see Table 1). Such tasks were used to allow for a direct comparison of search strategies and query formulations between the two groups. Imposed queries may be less motivating to solve than self-generated information needs (Gross, 1995). It was assumed that the participants were motivated to solve the tasks since they had volunteered for the study. Experimental task orders are often randomized to counterbalance learning effects. In this study, however, the search tasks were presented in a fixed order. The students had prior experience with the search system and it was assumed that the learning-effects would be minimal. The key design consideration was not to make the participants feel uncomfortable or lose courage during the session because of the varying difficulty levels of the tasks and poor self-esteem issues among many dyslexic students (Caroll and Iles, 2006). Consequently, the tasks were presented in increasing order of difficulty.

Task	Task (originally given in Norwegian)	Expected query terms	Translated query terms
1	Find a document about Vigeland Sculpture Park?	Vigelandsparken	Vigeland Sculpture Park
2	In which year was the book "Rock carvings in Hedmark and Oppland" published?	helleristninger Hedmark Oppland	rock carvings Hedmark Oppland
3	Find a book written by Sigrid Undset about Kristin?	Sigrid Undset Kristin	Sigrid Undset Kristin
4	Find a document on stagecraft and lighting?	sceneteknikk lyssetting	stagecraft lightning
5	Find a document on cyberbullying?	digital mobbing	cyberbullying
6	Who is the author of the books about Albert and Skybert?	Albert Skybert	Albert Skybert
7	Find a document about Norwegian recipients of the Nobel Peace Prize?	norske vinnere Nobels fredspris	Norwegian recipients Nobel Peace Prize
8	Find a book on Knut Hamsun and Nazism?	Knut Hamsun Nazisme	Knut Hamsun Nazism
9	Find a document about women in Algeria?	kvinner Algerie	women Algeria
10	Find a play written by William Shakespeare?	William Shakespeare	William Shakespeare

Table 1: Search tasks and expected query terms

The search tasks were based on gender-neutral and common topics. They were not related to any particular study programmes, preventing prior knowledge to affect search behaviour. One correct answer was deemed sufficient to solve each task. The tasks were designed to provoke spelling errors allowing for a study of how search behaviour is affected by reduced spelling skills. The tasks included words which dyslexic users frequently misspell, such as compound words: fredspris, double consonants: Oppland, consonant clusters: Undset, silent consonants: Sigrid (d is not pronounced), silent vowels: Algerie (e is not pronounced) and words with irregular orthography, where the spelling and pronunciation differs: Shakespeare (Helland and Kaasa, 2005; Lyster, 2007; Høien and Lundberg, 2012). Although the search tasks represent a moderate complexity level, it was assumed that the tasks would be perceived most difficult for dyslexic users.

Results

The query characteristics are summarised in Table 2. Dyslexic students took significantly longer to complete each task (*M*=47.8 s, *SD*=14.6 s) compared to controls (*M*=29.0 s, *SD*=8.2), $t(38)=5.02$, $p<.001$, $d=1.63$. Although the dyslexia group took longer on each query (*M*=19.8 s, *SD*= 5.1) than the control group (*M*=17.0 s, *SD*=3.6), the difference was not significant, $t(38)=1.95$, $p>.058$, $d=0.63$.

	Dyslexia group		Control group		
	Mean	SD	Mean	SD	p t-test)
Search time (s) per task	47.8	14.6	29.0	8.2	<.001
Search time (s) per query	19.8	5.1	17.0	3.6	>.058
Number of queries per task, overall	2.5	0.6	1.7	0.3	<.001
Number of queries per task, Bibsys	2.1	0.5	1.6	0.3	<.001
Number of queries per task, external	0.4	0.3	0.1	0.1	<.001
Portion of misspelled queries, Bibsys	44%	10%	26%	14%	<.001

Table 2: Query summary, SD denote standard deviations

Dyslexic users formulated significantly more queries per task (*M*=2.5, *SD*=0.6) than the controls (*M*=1.7, *SD*=0.3), $t(38)=5.42$, $p<.001$, $d=1.76$. A higher submission of queries was found both in the library catalogue and external Websites. In Bibsys Ask, users with dyslexia submitted significantly more queries per task (*M*=2.1, *SD*=0.5) than the controls (*M*=1. 6, *SD*=0.3), $t(38)=4.52$, $p<.001$, $d=1.47$. Furthermore, the dyslexic students used external Websites significantly more per task (*M*=0.4, *SD*=0.3) than the control group (*M*=0.1, *SD*=0.1), $t(38)=4.31$. $p<.001$, $d=1.40$. All external searches were conducted in Google, except one online dictionary query. No users in the control group relied on external Websites for task five, six and seven.

Dyslexic students misspelled more queries (*M*=0.44, *SD*=0.10) than the controls (*M*=0.26, *SD*=0.14), $t(38)=4.59$, $p<.001$, $d=1.49$. The number of spelling errors was higher for all search tasks. None of the controls made spelling errors on task five and six.

Concerning query lengths (see Table 3), dyslexic users included significantly fewer search terms per query (*M*=2.05, *SD*=0.23) than the controls (*M*=2.25, *SD*=0.18), $t(38)=2.92$, $p<.007$, $d=0.95$. However, when the first query was analysed separately from the remaining queries, both groups submitted similar query lengths, $t(38)=1.09$, $p>.281$, $d=0.35$. There were no significant differences in lengths in the last query for each task, $t(38)=1.87$, $p>.069$, $d=0.47$. Consequently, the shorter dyslexic query lengths appeared mainly in the intermediate queries.

	Dyslexia group		Control group		
	Mean	SD	Mean	SD	p t-test)
Submitted terms per query, overall	2.05	0.23	2.25	0.18	<.007
Submitted terms per first query	2.07	0.28	2.16	0.27	>.281
Submitted terms per last query	2.18	0.23	2.30	0.21	>.069
Minimum number of terms per query	1.75	0.33	2.01	0.20	<.006
Maximum number of terms per query	2.35	0.16	2.33	0.22	>.743

Table 3: Query lengths, SD denote standard deviations

The minimum number of terms per query was significantly lower among dyslexic users (M=1.75 terms, SD=0.33) than the controls (M=2.01 terms, SD=0.20), $t(38)=3.02$, $p<.006$, $d=0.98$. There were no significant differences in maximum query lengths, $t(38)=0.33$, $p>.743$, $d=0.10$.

The participants applied different approaches when a query failed to return useful results, for instance using external Websites or query modifications such as altering query lengths, changing the spelling of words or replacing words (see Table 4 and Figure 4).

	Dyslexia group		Control group		
	Mean	SD	Mean	SD	p t-test)
Portion of queries with changed spelling	46%	13%	40%	20%	>.252
Portion of queries where external web sites are applied	27%	11%	16%	15%	<.017
Portion of queries with increased query lengths	16%	8%	27%	24%	>.057
Portion of queries with decreased query lengths	11%	9%	10%	12%	>.774
Portion of queries with replacement of words	6%	9%	7%	16%	>.905

Table 4: Experimental results, SD denote standard deviations

The most frequent approach involved changing the spelling of terms. Dyslexic users changed the spelling in more queries (M=0.46, SD=0.13) than the controls (M=0.40, SD=0.20). However, there were no significant differences, $t(38)=1.16$, $p>.252$, $d=0.38$. Consulting external Websites to identify the correct spelling of terms was the only approach with significant differences. External Websites were used significantly more by dyslexics (M=0.27, SD=0.11) than controls (M=0.16, SD=0.15), $t(38)=2.53$, $p<.017$, $d=0.82$. There were no significant differences in the remaining approaches such as modifying query lengths or replacing terms (see Table 4).

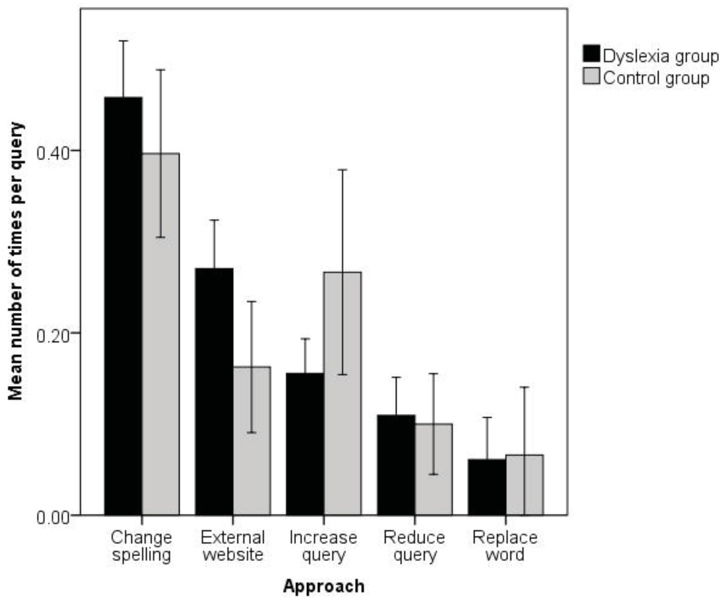


Figure 4: Mean use of query modification approaches per intermediate query

Discussion

Dyslexics often exhibit slow or inaccurate reading and writing (Caroll and Iles, 2006). It was therefore expected that the dyslexic users would take more time on each task than controls. This assumption was confirmed and is in accordance with the findings by MacFarlane et al. (2010). The increased time use could be explained by a higher number of queries and the use of external Websites. It is common that users with dyslexia have difficulties concentrating on a task for a long time (Mortimore and Crozier, 2006). A fatigue-related effect could therefore have caused increased search times. However, the time use did not increase during the experiment. Consequently, a more accessible interface may reduce search times for all tasks.

Users with dyslexia often have reduced rapid naming skills (Smith-Spark and Fisk, 2007), smaller vocabularies, are more limited in their choice of words (Connelly, Campbell, MacLean and Barnes, 2006) and display slower writing speeds than is usual for their age (Hatcher, Snowling and Griffiths, 2002). It was therefore expected that dyslexic users would take longer time per query compared to controls (H1), an assumption which was not confirmed. This result was unexpected. However, some of the search terms, such as author names, were included in the oral instructions, thus reducing the need for good rapid naming skills to quickly identify search terms. Moreover, dyslexic users formulated more short queries compared to the controls. Consequently, dyslexic users may have used more time per query if the query lengths had been similar throughout the experiment. These results imply that help functions which reduce the demands on query content may be helpful.

It was hypothesised that dyslexic users would formulate more queries per task than controls, an assumption which was confirmed. The observations indicate that the higher number of queries among the dyslexics was mainly caused by insufficient or too extensive result lists and misspellings, which led to query modifications and use of external Websites. These findings indicate that dyslexia has a negative effect on search performance in databases with a low tolerance for misspellings and no query-building aids.

Dyslexic users were expected to submit shorter queries than controls (H3). It was assumed that slower writing speed and a high risk of spelling errors would result in shorter queries, thus reducing search times and eliminating misspellings. Dyslexic students did submit significantly shorter queries than the controls. This difference was mainly found in the intermediate queries. There were no significant differences in query lengths when the first or last queries were analysed separately. Consequently, the participants started and ended the search sessions at approximately the same precision levels, which implies that the result lists contained approximately the same number of documents on a similar relevance level for both groups.

Query lengths affect the size and content of result lists. Precision and recall refers to a search engine's ability to retrieve relevant documents. There is a trade-off between these two measurements. Short and general queries result in long imprecise result lists, while longer and more specific queries lead to higher precision. Precision could potentially be an aim for slow readers. Precise queries may facilitate the browsing process by reducing the size of result lists and increasing the relevance of the first documents in the list. The costs of longer queries include a higher risk of misspellings, which may explain the shorter queries in the dyslexia group. Several dyslexic participants used the built-in search-function in the browser rather than scan the result lists. This behaviour suggests an attempt to reduce work load and solve the issues regarding the contradictory relationship between precision and recall. Moreover, these findings imply that there may be a need for alterations in the presentation of result lists to better accommodate users with dyslexia.

There was a significant difference between the groups regarding minimum query lengths, but not maximum lengths. Dyslexic students submitted shorter intermediate queries, possibly to isolate the misspelling, for instance:

Q_{1-1}: «knut hamsun nasismen», Q_{1-2}: «knut hamsun», Q_{1-3}: «knut hamsun nazismen».

The user removed one word in Q_{1-2}, to verify the spelling of the author name, before the subject term was reintroduced with a different spelling. Query Q2 is an example of a similiar approach, but here the user searched with one term at a time before combining them in the last query:

Q_{2-1}: «Sigrid Unseth», Q_{2-2}: «Unset», Q_{2-3}: «Kristin», Q_{2-4}: «Kristin Unset».

Query Q1 and Q2 indicate that a lack of query-building aids or a low tolerance for spelling errors may lead to more queries to solve a task. It would have been useful if the system provided the user with more substantial information when errors occurred, for instance by indicating which terms were not matched in the database: «*no results in the database for the term Unseth*» as a response to Q_{2-1}. This function would have counteracted the need for Q_{2-2} and Q_{2-3}, and thus cut the number of queries by a half.

It was assumed that the dyslexic students would exhibit more spelling errors than the controls (H4). This hypothesis was supported. The number of misspellings increased according to number of search terms, which was also as expected, since the use of more terms is likely to increase the total probability of spelling errors.

The number of errors was mainly related to query lengths and the inclusion of unfamiliar or difficult words, such as *helleristninger* and *Algerie*. Previous studies have found that adults with dyslexia may memorise the spelling of familiar words, but have more difficulties with orthographic patterns which results in more errors in unfamiliar words (Kemp, Parrila and Kirby, 2009). The results from the Bibsys searches are in accordance with these findings. Moreover, these findings also support the position that autocomplete functions could be beneficial when task complexity increases. This is in accordance with the students' comments during the de-briefing interview about the potential problematic features or shortcomings in search systems such as Bibsys Ask. All except one of the students with dyslexia mentioned the lack of spelling aids as the most prominent problem.

Although Bibsys Ask did not offer any query-building aids, the system included an authority file. Such files contain the authorised form of, for instance, author names as well as the addition of different spellings of terms. In these cases, users may retrieve results although a term is not spelled correctly if they have used a spelling which is in accordance with the variations in the file. It was anticipated that there would be many spelling errors in the William Shakespeare task. This task included terms with several challenging linguistic components, such as irregular orthography, double consonants and silent vowels. These complex linguistic components are typically difficult by dyslexic users. The last name of the author may also be regarded as an unfamiliar word. However, the library catalogue actually returned results when the query was formulated without the 'e' at the end of Shakespeare, because of the authority file. The participants also received results when adding a 'd' at the end of Hamsun. If such a mapping had not been included, the number of registered misspellings and query modifications would presumably have been even higher. Further, the use of external Websites was higher for the Shakespeare task, thus reducing the number of misspelled queries.

Several approaches for resolving misspellings were observed. Nevertheless, many errors could have been averted through autocomplete functions or spelling suggestions. However, in many searches a higher tolerance for spelling errors would have been adequate to accommodate dyslexic users. This was particularly evident in the number of Google searches where relevant results were retrieved despite misspelled queries. These results indicate that a high tolerance for spelling errors reduce the negative effect of dyslexia on search performance. Users with dyslexia often make different spelling errors than their peers (Li, Sbatella and Tedesco, 2013). Consequently, spellcheckers which accommodate the misspellings of users with dyslexia should be implemented. Such technology is already available in proprietary software. Participants without dyslexia also made spelling errors during the search sessions which shows that query-building aids would be useful for a wide cohort of users. This finding supports the work of de Santana et al. (2012).

Many dyslexics in the study claimed that they often guessed the spelling of words rather than spending time compiling the correct query. In other cases Google was used as a dictionary. Schoot, Licht, Horsley and Sergeant (2000) found that guessing while reading is not uncommon for dyslexics. However, results from this Bibsys Ask study may indicate that such guessing behaviour may also be applied during writing.

Users with dyslexia often develop coping strategies to compensate for their impairment (Borkowski, Estrada, Milstead and Hale, 1989). Consequently, it was hypothesised that dyslexic users would apply a broader range of problem solving strategies than controls (H5). However, using external Websites was the only approach that was significantly undertaken by the dyslexic users.

Google was the only external Website used by participants in the study, except for one direct search in an online dictionary. This was not surprising, since Google is a widely used search engine (Westerwick, 2013), and it has a high tolerance for spelling errors and includes autocomplete functions. The high tolerance level for misspelled queries was evident in several searches in this study. For instance, a dyslexic student received relevant results on a query despite three misspellings: «*helerystninger i hedemark og oppland*» (correct query «*helleristninger i hedmark og oppland*»).

Using external resources may be an approach based on prior experiences. Such behaviour may also be an example of least possible effort (Bates, 2002) where students try to avoid the system most difficult to use, since Google demands less effort from the users. A majority of the students found the library catalogues difficult to use compared to general search engines. Consequently, they preferred formulating the query in Google to avoid using the library catalogue.

Several dyslexics usually searched Google first to find the proper spelling before pasting the query into the library catalogue search interface. Such a strategy reduced the effort and mental load required in Bibsys Ask to compile the proper query syntax. This approach was apparent in searches with difficult spellings, where students in both groups searched Google directly on the first query. Several students mentioned that such behaviour was a result of past experiences with faulty queries in Bibsys Ask. If the students had not applied this strategy, the number of queries and misspellings could have been even more prevailing. Several students with dyslexia often used Google instead of actual online dictionaries for spell checking of Norwegian words in addition to Google translate for English terms. This may indicate that users regard Google as more accessible compared to online reference works.

Some participants revealed that they normally used Google more extensively than during this particular study. In the experimental setting they wanted to try to solve the tasks in Bibsys Ask as instructed. Consequently, the use of external resources may be even more common in the participants' usual practices.

Queries submitted to Google were often longer and less formal than those submitted to Bibsys Ask. For instance, one student submitted the query «*nobelsfredspris*» (correct term *Nobels fredspris*,) in Bibsys Ask, which returned zero results. The student then submitted the following query to Google: «*Norske vinnere av nob*», whereas Google's autocomplete suggested a correct query. This query was copied and pasted into Bibsys Ask. The increased query lengths in Google may be a response to a high tolerance for spelling errors or that the autocomplete function suggests longer queries. In this study, participants uncertain about the spelling of terms typically inputted a few letters in Google's search interface and then selected or looked at the correct spelling in the autocomplete list. For instance, one user inputted «*nob*» on the task on Norwegian Nobel Peace Prize recipients and «*knut*» on the Knut Hamsun and Nazism task. These findings suggest that autocomplete may be helpful. Moreover, Google offers full text searching. Consequently, it may be more useful to submit longer informal queries than in bibliographic databases such as Bibsys Ask, which may explain also explain increased query lengths.

A total of 19 dyslexic users and 13 controls searched Google after submitting an erroneous query in Bibsys Ask, rather than modifying the query in the library catalogue. However, several participants actually resolved the misspellings themselves through the altered query submitted to Google. Consequently, the students could have solved the task directly in Bibsys Ask without relying on Google. This may imply that users do not want to spend much time troubleshooting, but rather go directly to a more familiar and tolerant search system when errors occur. Such behaviour may also be caused by a lack of confidence, which is often seen in users with dyslexia (Jordan, McGladdery and Dyer, 2014).

In one case where Google returned results despite misspellings the user was not given feedback of the spelling error. When a relevant result list was presented despite the misspelling, the student did not realise the spelling error and continued searching Bibsys Ask with the erroneous query. Consequently, this user took more than three minutes solving the task. This behaviour indicates that users may rely on feedback of mistakes. This is also in accordance with general principles of human computer interaction design (Norman, 2013).

Users seem to rely quite uncritically on the suggestions provided by Google. In two searches for Norwegian recipients of the Nobel Peace Prize, the dyslexic users got the suggestion «*norske nobel vinnere*» by Google, which was pasted into Bibsys Ask. This was not an incorrect suggestion, since these two words do have meaning separately in Norweian. However, in this particular task, the last two words had to be one compound word: «*norske nobelvinnere*» to solve the task. The users did not notice that the query was incorrect and submitted three additional queries each to Bibsys Ask. This finding indicates that Google´s query-building tool sometimes is suboptimal.

Changing spelling was the most frequent problem solving approach within both groups. The dyslexic users seemed especially aware of common mistakes such as double consonants and compound words, and rephrased queries accordingly, for instance:

Q_{3-1}: «*scene teknikk*», Q_{3-2}: «*sceneteknikk*».

Another user tried different ways of spelling, based on double consonants, in addition to the nd- and dt-sound which may also be problematic for users with dyslexia:

Q_{4-1}: «*Sigrid Undsett*», Q_{4-2}: «*Sigrid Unsett*», Q_{4-3}: «*Sigrid Unnsett*», Q_{4-4}: «*Sigrid Unnsedt*» (...) (correct spelling Sigrid Undset)

These queries indicate that the search performance would have been approved if the system had suggested the correct spelling.

Changing the spelling of a term is a self-evident strategy for modifying a misspelled query. However, other approaches included replacing or excluding difficult words. Such strategies were especially noticeable in the task on the author of the books about Albert and Skybert (in Norwegian the Sk is pronounced sh, making it potentially difficult to spell). On this task, 25% of the dyslexic users searched with «*Albert Åberg*». This query was based on prior knowledge of this children's book character's last name, which was not included in the instructions, thereby avoiding a difficult term. This strategy was also found in 25% of the searches by the controls. The results from this study indicate that word replacement might be a common strategy for all users, and is not a particular characteristic of dyslexic users.

Replacements were also applied in intermediate queries. For instance, one student with dyslexia used this approach when looking for «*Helleristninger i Hedmark og Oppland*» and Hedmark was regarded difficult to spell ($Q_{5\text{-}4}$):

$Q_{5\text{-}1}$: «*helleristinger*», $Q_{5\text{-}2}$: «*helleristinger hedmark*», $Q_{5\text{-}3}$: «*helleristinger hedemark*», $Q_{5\text{-}4}$: «*helleristinger i oppland*»

The replacement approach may not always be successful in bibliographic databases such as Bibsys Ask. In a full text database it is plausible that synonyms are applied to vary the writing style. However, in a bibliographic database the query should match the terms included in the metadata, such as names, titles or subject headings. However, tools such as synonym control or thesauri may provide results although the query contains non-preferred terms. In these cases the replacement approach could be productive. Consequently, search systems should support such functions to accommodate users who exhibit this type of query modification.

Finally, several students with dyslexia mentioned asking a librarian as a commonly used problem solving approach. One dyslexic student had given up searching the library catalogue altogether, and always got help from librarians because locating literature by using the database was too time-consuming. This is in accordance with Mortimore and Crozier's (2006) findings that a majority of the students with dyslexia in higher education used or wanted extra library support. It is also an example of least possible effort (Bates, 2002). However, such situations should be avoided by providing all users with accessible interfaces and excessive training.

Conclusion

The results from this study indicate that users with dyslexia struggle when searching databases with a low tolerance for errors and no query-building aids. Dyslexic users took more time, formulated more queries, submitted shorter queries and made more misspellings than controls. Further, they relied more on external Websites to solve tasks. However, if functions that reduce the demands for correct spelling are implemented, the interface may become more accessible and thus reduce the negative effect of dyslexia. For instance, search systems could contain the following functions:

- feedback on which part of a query where there is no match
- suggest terms if misspellings occur
- allow users to replace difficult terms (synonym control)
- suggest queries while the user is entering queries (autocomplete)
- tolerance for spelling errors, including errors typically made by dyslexic users

Moreover, the result lists should be more accessible, for instance by including more visual information such as images or icons. Further research is needed to identify more effective result list designs and investigate whether textual or visual content is easiest to browse through for dyslexic users.

Users without reading or writing disorders also make spelling errors, which was evident in this study. Misspellings can occur for variety of reasons, for instance if users are inputting text too fast, using an unfamiliar keyboard, inputting unknown words or if the user is tired or unfocused. Consequently, improving an interface to accommodate dyslexic users will enhance the search experience for all users. This finding is consistent with other studies (McCarthy and Swierenga, 2010; Boldyreff, Burd, Donkin and Marshall, 2001).

No significant differences in problem solving approaches were uncovered, except for a more frequent use of external resources among dyslexic users. The widespread use of a system with query-building aids and a high tolerance for errors may indicate that such systems are perceived more accessible. However, this can also be attributable to more experience with Google than the library catalogue.

Query lengths indicated that there were no significant differences in precision levels, which implies that users with and without dyslexia may have similar basic search skills. The differences in search behaviour may be merely caused by impaired reading and writing skills. However, since many search terms were included in the assignments, the study did not reveal if dyslexia affected the selection of proper search terms.

Results from this study seem to strengthen the assumption that search interfaces with a low tolerance for spelling errors and no query-building aids are not accessible for users with dyslexia. More specific guidelines are needed to reduce the need for extra support and strengthen the independence, thus reducing the need for extra library support.

Acknowledgements

This project has been financially supported by the Norwegian ExtraFoundation for Health and Rehabilitation through EXTRA funds grant 2011/12/0258. The authors are grateful to the volunteering students and Dyslexia Norway for help recruiting participants.

About the authors

Gerd Berget is a PhD student at Institute of Technology at Oslo and Akershus University College of Applied Sciences. Her research includes universal design, information searching behavior, dyslexia and eye tracking studies. She can be contacted at gerd.berget@hioa.no.
Frode Eika Sandnes is a professor at the Institute of Technology at Oslo and Akershus University College of Applied Sciences and at the Faculty of Technology, Westerdals Oslo School of Arts, Communication and Technology. Oslo, Norway. His research interests include human computer interaction, universal design and pattern matching. He can be contacted at frode-eika.sandnes@hioa.no.

References

Al-Maskari, A. & Sanderson, M. (2011). The effect of user characteristics on search effectiveness in information retrieval. *Information Processing & Management*, *47*(5), 719-729. Retrieved from http://bit.ly/1RfhLYL. (Archived by WebCite®® at http://www.webcitation.org/6a0gsS7Qf)

Allen, D. K. & Wilson, T. D. (2003). Information overload: context and causes. *New Review of Information Behaviour Research*, *4*(1), 31-44

Al-Wabil, A., Zaphiris, P. & Wilson, S. (2007). Web navigation for individuals with dyslexia: an exploratory study. In C. Stephanidis (Ed.), *Universal access in human computer interaction:coping with diversity*. (pp. 593-602). Berlin: Springer. (Lecture Notes in Computer Science Volume 4554). Retrieved from http://openaccess.city.ac.uk/1051/1/1406_final.pdf. (Archived by WebCite® at http://www.webcitation.org/6a0h3PBaS)

Association of College and Research Libraries (2000). *Information literacy competency standards for higher education*. Chicago, IL: American Library Association. Retrieved from https://arizona.openrepository.com/arizona/bitstream/10150/105645/1/standards.pdf. (Archived by WebCite®® at http://www.webcitation.org/6a0h8QOzr)

Badian, N. A. (1999). Reading disability defined as a discrepancy between listening and reading comprehension: a longitudinal study of stability, gender differences, and prevalence. *Journal of Learning Disabilitiesm*, *32*(2), 138-148.

Bates, M. J. (2002). Toward an integrated model of information seeking and searching. *New Review of Information Behaviour Research*, **3**, 1-15. Retrieved from http://ptarpp2.uitm.edu.my/silibus/TOWARDANINTEGRATEDMODEL.pdf. (Archived by WebCite® at http://www.webcitation.org/6a0jAc9rZ)

Berget, G. & Sandnes, F. E. (in press). Do autocomplete functions reduce the effect of dyslexia on information searching behavior? The case of Google. *Journal of the American Society for Information Science & Technology*, in press.

Boldyreff, C., Burd, E., Donkin, J. & Marshall, S. (2001). The case for the use of plain English to increase Web accessibility. In *Proceedings of the 3rd International Workshop on Web Site Evolution, Florence, Italy*. (pp. 42-48). Washington, DC: IEEE.

Borkowski, J. G., Estrada, M. T., Milstead, M. & Hale, C. A. (1989). General problem-solving skills: relations between metacognition and strategic processing. *Learning Disability Quarterly, 12*(1), 57-70.

Caroll, J. M. & Iles, J. E. (2006). An assessment of anxiety levels in dyslexic students in higher education. *British Journal of Educational Psychology, 76*(3), 651-662. Retrieved from http://wrap.warwick.ac.uk/385/1/WRAP_Carroll_anxietyJMC_140305.pdf. (Archived by WebCite® at http://www.webcitation.org/6a0jFJdsk)

Connelly, V., Campbell, S., MacLean, M. & Barnes, J. (2006). Contribution of lower order skills to the written composition of college students with and without dyslexia. *Developmental Neuropsychology, 29*(1), 175-196. Retrieved from http://bit.ly/1GoKBlr. (Archived by WebCite® at http://www.webcitation.org/6a0jKkAHz)

Craven, J. & Brophy, P. (2003). *Non-visual access to the digital library: the use of digital library interfaces by blind and visually impaired people*. Manchester, UK: Manchester Metropolitan University, Centre for Research in Library and Information Management.

Dahle, A. E. & Knivsberg, A.-M. (2014). Internalizing, externalizing and attention problems in dyslexia. *Scandinavian Journal of Disability Research, 16*(2), 179-193.

de Santana, V. F., de Oliveira, R., Almeida, L.D.A., & Baranauskas, M. C. C. (2012). Web accessibility and people with dyslexia: a survey on techniques and guidelines. In M. Vigo, J. Abascal, P. Salomoni and R. Lopes (Eds.), *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility (W4A), Lyon, France, 16-17 April, 2012*. New York, NY: ACM. Retrieved from http://www.ramb.ethz.ch/CDstore/www2012/W4A/01%20-%20Technical%20Papers/santana.pdf. (Archived by WebCite® at http://www.webcitation.org/6a0jPMH9d)

Evett, L. & 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.

Freire, A.P., Petrie, H. & Power, C. (2011). Empirical results from an evaluation of the accessibility of websites by dyslexic users. In G. Weber, H. Petrie and J. Darzentas (Eds.), *Proceedings of the Workshop on Accessible Design in the Digital World 2011, CEUR, Lisbon, Portugal, September 5, 2011*, (pp. 41-53). Aachen, Germany: CEUR-WS.Org. Retrieved from https://www.cs.york.ac.uk/hci/publications/002/dyslexia_addw%20v4.pdf. (Archived by WebCite® at http://www.webcitation.org/6a0jThrTg)

Germanó, E., Gagliano, A. & Curatolo, P. (2010). Comorbidity of ADHD and dyslexia. *Developmental Neuropsychology, 35*(5), 475-493.

Gross, M. (1995). The imposed query. *Reference & User Services Quarterly, 35*(2), 236–243.

Habib, L., Berget, G., Sandnes, F.E., Sanderson, N., Kahn, P., Fagernes, S. & Olcay, A. (2012). Dyslexic students in higher education and virtual learning environments: an exploratory study. *Journal of Computer Assisted Learning, 28*(6), 574-584.

Hatcher, J., Snowling, M.J. & Griffiths, Y.M. (2002). Cognitive assessment of dyslexic students in higher education. *British Journal of Educational Psychology, 72*(1), 119-133.

Hawke, J.L., Olson, R.K., Willcut, E.G., Wadsworth, S.J. & DeFries, J.C. (2009). Gender ratios for reading difficulties. *Dyslexia, 15*(3), 239-242.

Helland, T. & Kaasa, R. (2005). Dyslexia in English as a second language. *Dyslexia, 11*(1), 41-60.

Henry, S. L., Abou-Zahra, S. & Brewer, J. (2014). The role of accessibility in a universal web. In J. P. Bigham, Y. Borodin and Carriço (Eds.), *W4A '14, April 07 - 09 2014, Seoul, Republic of Korea*, ACM, New York, article no. 17. Retrieved from http://dspace.mit.edu/openaccess-disseminate/1721.1/88013. (Archived by WebCite® at http://www.webcitation.org/6a0jYe1o8)

Hepworth, M. (2007). Knowledge of information behaviour and its relevance to the design of people-centred information products and services. *Journal of Documentation, 63*(1), 33–56.

Høien, T. & Lundberg, I. (2012). *Dysleksi: fra teori til praksis* [Dyslexia: from theory to practise]. Oslo: Gyldendal.

Høien, T. & Tønnesen, G. (2008). *Ordkjedetesten* [The word chain test]. Bryne, Norway: Logometrica.

Imrie, R. (2012). Universalism, universal design and equitable access to the built environment. *Disability and Rehabilitation, 34*(10), 873-882.

ISO. (2009). *ISO 8596 Ophthalmic optics: visual acuity testing – standard optotype and its presentation*. Geneva, Switzerland: International Organization for Standardization.

Jamali, H.R., & Asadi, S. (2010). Google and the scholar: the role of Google in scientists' information-seeking behaviour. *Online Information Review, 34*(2), 282–294. Retrieved from http://bit.ly/1RvtlLH. (Archived by WebCite® at http://www.webcitation.org/6a0jdzuBc)

Jones, M.W., Branigan, H.P. & Kelly, M.L. (2007). Visual deficits in developmental dyslexia: relationships between non-linguistic visual tasks and their contribution to components of reading. *Dyslexia, 14*(2), 95-115.

Jordan, J-A., McGladdery, G. & Dyer, K. (2014). Dyslexia in higher education: implications for maths anxiety, statistics anxiety and psychological well-being. *Dyslexia, 20*(3), 225-240.

Kemp, N., Parrila, R. K. & Kirby, J. R. (2009). Phonological and orthographic spelling in high-functioning adult dyslexics. *Dyslexia, 15*(2), 105-128. Retrieved from http://bit.ly/2062ZpM. (Archived by WebCite® at http://www.webcitation.org/6a0jiNYK5)

Kim, K-S. & Allen, B. (2002). Cognitive and task influences on web searching behaviour. *Journal of the American Society for Information Science & Technology, 53*(2), 109-119.

Kuhlthau, C. C. (1991). Inside the search process: information seeking from the user's perspective. *Journal of the American Society for Information Science, 42*(5), 361-371. Retrieved from http://ptarpp2.uitm.edu.my/silibus/insidesearch2.pdf. (Archived by WebCite® at http://www.webcitation.org/6a0jmjiLb)

Lervåg, A., & Hulme, C. (2009). Rapid automatized naming (RAN) taps a mechanism that places constraints on the development of early reading fluency. *Psychological Science, 20*(8), 1040–1048. Retrieved from http://thinkingskillsclub.com/wp-content/uploads/2014/04/ran_brain_area_2009-2010.pdf. (Archived by WebCite®

at http://www.webcitation.org/6a0juy6j9)

Li, A.Q., Sbatella, L. & Tedesco, R. (2013). PoliSpell: an adaptive spellchecker and predictor for people with dyslexia. In S. Carberry, S. Weibelzahl, A. Micarelli & G. Semeraro (Eds.), *UMAP 2013, The 21st Conference on User Modeling, Adaptation and Personalization, Rome, Italy, Springer, Berlin*, pp. 302-309.

Lyster, S-A.H. (2007). Reading development and reading disabilities: focus on Norway. In B. G. Cook, M. Tankersley, & T. J. Landrum (Eds.) *Advances in Learning and Behavioral Disabilities, 20*, 21-55.

MacFarlane, A., Al-Wabil, A., Marshall, C. R, Albrair, A., Jones, S. A. & Zaphiris, P. (2010). The effect of dyslexia on information retrieval: a pilot study. *Journal of Documentation, 66*(3), 307-326. Retrieved from http://openaccess.city.ac.uk/1694/1/dyslexia-paper-JDOC-revised-FINAL.pdf. (Archived by WebCite® at http://www.webcitation.org/6a0kC0kHm)

MacFarlane, A., Albrair, A., Marshall, C.R. & Buchanan, G. (2012). Phonological working memory impacts on information searching: an investigation of dyslexia. In *Proceedings of IIiX 2012 the Fourth Information Interaction in Context Symposium, Nijmegen, the Netherlands, August 21-24, 2012*. (pp. 27-34). New York, NY: ACM. Retrieved from http://openaccess.city.ac.uk/4493/1/dyslexia-memory-search-oa.pdf. (Archived by WebCite® at http://www.webcitation.org/6a0kGrW5k)

McCarthy, J.E. & Swierenga, S.J. (2010). What we know about dyslexia and web accessibility: a research review. *Universal Access in the Information Society, 9*(2), 147-152. Retrieved from http://bit.ly/1WccYu4. (Archived by WebCite® at http://www.webcitation.org/6a0kL1HMT)

Mortimore, T. & Crozier, W. R. (2006). Dyslexia and difficulties with study skills in higher education. *Studies in Higher Education, 31*(2), 235-251.

Norman, D. (2013). *The design of everyday things*. New York, NY: Basis Books.

Palmquist, R. A. & Kim, K.-S. (2000). Cognitive style and on-line database search experience as predictors of web search performance. *Journal of the American Society for Information Science & Technology, 51*(6), 558-566.

Papaeconomou, C., Zijlema, A. F., & Ingwersen, P. (2008). Searchers' relevance judgments and criteria in evaluating web pages in a learning style perspective. In P. Borlund, J. W. Schneider, M. Lamas, A. Tombros, J. Feather, D. Kelly & A. P. de Vries (Eds.) *Proceedings of the second international symposium on information interaction in context, London, October 14-17*, (pp 123–132). New York, NY: ACM.

Ransby, M. J. & Swanson, H. L. (2003). Reading comprehension skills of young adults with childhood diagnoses of dyslexia. *Journal of Learning Disabilities, 36*(6), 538-555.

Rello, L. & Baeza-Yates, R. (2013). Good fonts for dyslexia. In *ASSETS '13, The 15th International ACM SIGACCESS Conference on Computers and Accessibility, Bellevue, WA, USA, October 21 – 23*. (Article No. 4). New York, NY: ACM

Rello, L., Pielot, M., Marcos, M-C. & Carlini, R. (2013). Size matters (spacing not): 18 points for a dyslexic-friendly Wikipedia. In *W4A'13, International Cross-Disciplinary Conference on Web Accessibility, May 13-15, 2013, Rio de Janeiro, Brazil*, (Article no. 17). New York, NY: ACM.

Rømen, D. & Svanæs, D. (2012). Validating WCAG versions 1.0 and 2.0 through usability testing with disabled users. *Universal Access in the Information Society, 11*(4), 375-385.

Sahib, N. G., Tombros, A. & Stockman, T. (2012). A comparative analysis of the information-seeking behavior of visually impaired and sighted searchers. *Journal of the American Society for Information Science, 63*(2), 377-391. Retrieved from http://ejournals.library.ualberta.ca/index.php/EBLIP/article/downloadSuppFile/18556/2757. (Archived by WebCite® at http://www.webcitation.org/6a0kpJfZw)

Schoot, M. v d, Licht, R., Horsley, T. M. & Sergeant, J. A. (2000). Inhibitory deficits in reading disability depend on subtype: guessers not spellers. *Child Neuropsychology, 6*(4), 297-312. Retrieved from http://dspace.ubvu.vu.nl/bitstream/handle/1871/18439/Schoot_Child%20Neuropsychology_6%284%29_2000_u.pdf?sequence=2. (Archived by WebCite® at http://www.webcitation.org/6a0kvAIP2)

Schuchardt, K., Maehler, C. & Hasselhorn, M. (2008). Working memory deficits in children with specific learning disorders. *Journal of Learning Disabilities, 41*(6), 514-523.

Smith-Spark, J.H. & Fisk, J.E. (2007). Working memory functioning in developmental dyslexia. *Memory, 15*(1), 34-56.

Snowling, M.J. (2000), *Dyslexia*. Malden, MA: Blackwell.

Snowling, M.J. (2001). From language to reading and dyslexia. *Dyslexia, 7*(1), 37-46.

Solheim, O. J. & Uppstad, P. H. (2011). Eye-tracking as a tool in process-oriented reading test validation. *International Electronic Journal of Elementary Education, 4*(1), 153-168. Retrieved from http://brage.bibsys.no/xmlui/bitstream/handle/11250/185914/Eye-tracking_as_a_tool.pdf?sequence=1. (Archived by WebCite® at http://www.webcitation.org/6a0l267NM)

Swanson, H.L., & Hsieh, C-J. (2009). Reading disabilities in adults: a selective meta-analysis of the literature. *Review of Educational Research, 79*(4), 1362–1390.

Tops, W., Callens, M., Lammertyn, J., Van Hees, V. & Brysbaert, M. (2012). Identifying students with dyslexia in higher education. *Annals of Dyslexia, 62*(3), 186-203.

Westerwick, A. (2013). Effects of sponsorship, web site design and, and Google ranking on the credibility of online information. *Journal of Computer-Mediated Communication, 18*(2), 194-211. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/jcc4.12006/full. (Archived by WebCite® at http://www.webcitation.org/6a0l9G3j5)

World Wide Web Consortium. (2008). *Web content accessibility guidelines (WCAG) 2.0*. Retrieved from http://www.w3.org/TR/WCAG20/. (Archived by WebCite® at http://www.webcitation.org/6cXw1tqwt)

Zhang, P., Bobier, W., Thompson, B. & Hess, R.F. (2011). Binocular balance in normal vision and its modulation by mean luminance. *Optometry and Vision Science, 88*(9), 1072-1079.

How to cite this paper

Berget, G. & Sandnes, F-E. (2015). Searching databases without query-building aids: implications for dyslexic users. *Information Research, 20*(4), paper 689. Retrieved from http://InformationR.net/ir/20-4/paper689.html (Archived by WebCite® at http://www.webcitation.org/6cdc2O2gH)

Find other papers on this subject

Scholar Search | Google Search | Bing

Check for citations, using Google Scholar

Contents | Author index | Subject index | Search | Home

Appendix D

Title:	Do autocomplete functions reduce the effect of dyslexia on information searching behaviour? The case of Google
Authors:	Gerd Berget and Frode Eika Sandnes
Journal:	Journal of the American Society for Information Science & Technology
Year:	Accepted 2015 (early-view published online October 2015)

Appendix E

Title: On the understandability of public domain icons: Effects of gender and age

Authors: Gerd Berget and Frode Eika Sandnes

Proceeding: M. Antona & C. Stephanidis (Eds.), Universal Access in Human-Computer Interaction: Access to Today's Technologies, 9th International Conference, UAHCI 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings, Part I

Publisher: Springer

Year: 2015

Pages: 387-396

Appendix F

Title: Search, read and write: An inquiry into Web accessibility for dyslexics

Authors: Gerd Berget, Jo Herstad and Frode Eika Sandnes

Proceeding: H. Petrie, J. Darzentas, T. Walsh, D. Swallow, L. Sandoval, A. Lewis & C. Power (Eds), Universal Design 2016: Learning from the Past, Designing for the Future

Publisher: IOC Press

Year: 2016

Pages: 450-460

Universal Design 2016: Learning from the Past, Designing for the Future
H. Petrie et al. (Eds.)

doi:10.3233/978-1-61499-684-2-450

Search, Read and Write: An Inquiry into Web Accessibility for People With Dyslexia

Gerd BERGET[ab,1], Jo HERSTAD[c] and Frode Eika SANDNES[bd]
[a] *Department of Archivistics, Library and Information Science, Oslo and Akershus University College of Applied Sciences, Oslo, Norway*
[b] *Institute of Information Technology, Oslo and Akershus University College of Applied Sciences, Oslo, Norway*
[c] *Research Group for Design of Information Systems, Faculty of Mathematics and Natural Sciences, University of Oslo, Norway*
[d] *Westerdals Oslo School of Art, Communication and Technology, Oslo, Norway*

Abstract. Universal design in context of digitalisation has become an integrated part of international conventions and national legislations. A goal is to make the Web accessible for people of different genders, ages, backgrounds, cultures and physical, sensory and cognitive abilities. Political demands for universally designed solutions have raised questions about how it is achieved in practice. Developers, designers and legislators have looked towards the Web Content Accessibility Guidelines (WCAG) for answers. WCAG 2.0 has become the de facto standard for universal design on the Web. Some of the guidelines are directed at the general population, while others are targeted at more specific user groups, such as the visually impaired or hearing impaired. Issues related to cognitive impairments such as dyslexia receive less attention, although dyslexia is prevalent in at least 5-10% of the population. Navigation and search are two common ways of using the Web. However, while navigation has received a fair amount of attention, search systems are not explicitly included, although search has become an important part of people's daily routines. This paper discusses WCAG in the context of dyslexia for the Web in general and search user interfaces specifically. Although certain guidelines address topics that affect dyslexia, WCAG does not seem to fully accommodate users with dyslexia.

Keywords. Universal Design, dyslexia, Web accessibility guidelines, WCAG, search

1. Introduction

Browsing and searching comprise the fundamental means of navigating Web content [1]. This behaviour is in accordance with the suggestions by Berners-Lee and Cailiau [2] when they introduced the Web, namely that searchable indexes should be added to the hypertext model. Search has become an integrated part of people's daily routines [3]. However, information retrieval may be difficult, especially for users who struggle

[1] Corresponding Author, Oslo and Akershus University College of Applied Sciences, Postboks 4 St. Olav plass, N-0130 Oslo, Norway; Email: gerd.berget@hioa.no.

with spelling and reading, such as people with dyslexia [4-5]. Consequently, accessible search user interfaces are necessary to ensure all users equal access to online content.

Universal design promotes equality and fairness, and entails that the society should be accessible for all users without the need for any adaptation or specialised design [6]. An important aspect of universal design is that users are not considered one homogenous group. Although most people share basic physiological and psychological characteristics, differences occur. Universal design incorporates variations in physical and cognitive abilities in addition to differences in gender, age and cultural background.

The UN Convention on the Rights of Persons with Disabilities states that information and communication technologies should be universally designed [7]. Universal design is also included in the EU eAccessibility programme [8] and in national legislations. A challenge related to legislation is *how* to define products or services as universally designed, and a lack of formal regulations may make it difficult to enforce the law. Consequently, WCAG has been applied in the legislation of several countries, such as Norway [9], Canada [10] and Australia [11].

According to Huffaker [12], WCAG is now the most recognised framework for accessibility rating. The guidelines are developed by the World Wide Web Consortium (W3C) and address the needs of a wide range of users. W3C [13] mentions for example blindness and low vision, deafness and hearing loss, learning disabilities, cognitive limitations, limited movement, speech disabilities, photosensitivity and a combination of these impairments. An overall aim is equal and barrier free access for all. Dyslexia is not mentioned specifically, but is a part of the broader terminology applied, namely «*cognitive, language and learning disabilities*» and «*print disabilities*» [13].

Accommodating a wide range of users may be difficult, for instance due to conflicting needs or because certain users may have difficulties expressing their needs [14]. This issue is also addressed in the WCAG introduction, which states: «*Note that even content that conforms at the highest level (AAA) will not be accessible to individuals with all types, degrees, or combinations of disability, particularly in the cognitive language and learning areas*». Dyslexia is an example of a common learning disorder which falls under the category described by W3C in their emphasis on potential shortcomings in WCAG. This cognitive impairment is mainly characterised by reading and writing difficulties [15] and is prevalent in 5-10% of the population [16].

It is claimed that the Web is not accessible for users with dyslexia, and that Web content may be inaccessible for people with dyslexia despite compliance with WCAG [17-18]. The Disability Rights Commission [19] reported that users with dyslexia failed on 17% of the tasks in a Web usability test. Two issues related to navigation, mainly unclear and confusing layout and confusing and disorienting navigation mechanisms. Two difficulties involved design, and were caused by insufficient contrast and too small graphics or text. The last issue involved complicated language and terminology.

Online search systems have been reported to be particularly inapproachable for people with dyslexia, especially systems with a low tolerance for spelling errors [5,17,20]. Inaccessible search systems represent a challenge in a digitalised society, where access to online information has become an integrated part.

This paper discusses WCAG in relation to dyslexia, and questions whether WCAG sufficiently accommodates people with dyslexia. Accessibility is discussed in the context of the Web in general, and with a particular focus on search user interfaces. The paper is structured as follows: First, related research on dyslexia and Web accessibility is presented, followed by a description of WCAG and how people with dyslexia are accommodated. Then, the guidelines are discussed according to dyslexia.

2. Background

Dyslexia exists in different forms and degrees. Some users with dyslexia are slow but accurate readers, while others are fast but inaccurate readers [21]. Nevertheless, difficulties with word decoding and reading comprehension are commonly found in both types of readers [22]. Other common characteristics are reduced short-term memory capacity, concentration difficulties and reduced naming skills [15, 23-25].

Dyslexia may affect the use of computers in several ways. Some struggle using keyboards due to reduced motor skills and coordination [26]. Frequent misspellings may cause difficulties using Web pages that require textual input from the user, such as search systems or other types of forms [5]. Slow reading, reduced decoding skills or reduced short term memory capacity can cause difficulties understanding textual content, such as menus or instructions and also affect result list assessment [27].

2.1 Web Accessibility and People with Dyslexia

Several researchers have addressed dyslexia and Web accessibility. Al-Wabil, Zaphiris & Wilson [17] interviewed ten people with dyslexia about their navigation behaviour. They revealed significant navigational barriers online. Kurniawan and Conroy [28] reported increased reading speed when users with dyslexia were allowed to select colour schemes. Rello, Kanvinde & Baeza-Yates [29] found that there was no coherence between preferences regarding colour schemes and performance. This finding is in accordance with Nielsen [30], who claims that users do not always prefer the solution in which they perform the best.

It has been suggested that people with dyslexia prefer fonts without serifs [31], and that italics should be avoided [32]. Font size may also have a significant effect on readability [31, 33]. Increased letter-spacing may improve reading performance [34], but line spacing has no significant effect [33]. Further, it is reported that users with dyslexia prefer left-justified text with a ragged right edge [35].

Graphics have also been discussed, for instance the accessibility of graphics in technical documentation [36] and how people with dyslexia interpret graphs [37]. Others have discussed whether including graphic content in text interfaces are useful [38-40]. The issue of clear and readable text in relation to users with dyslexia or low-proficiency readers has also been addressed. Rello, Baeza-Yates, Dempere-Marco and Saggion [41] found that using frequent words increased the reading speed of people with dyslexia, while shorter words enhanced the understanding of the content. The importance of using plain language is also suggested to increase Web accessibility for all users [42-43].

2.2 Accessible Search and People with Dyslexia

Information search is a common activity, and millions of searches are conducted on a daily basis [3]. Al-Wabil et al. [17] reported general difficulties with search among users with dyslexia, and specifically in internal search systems with high demands for spelling. The finding was supported by Habib et al. [20]. MacFarlane et al. [4] reported that students with dyslexia exhibited fewer search iterations, reviewed less documents and took longer per search than people without dyslexia. MacFarlane, Albrair, Marshall and Buchanan [27] investigated how impaired short-term memory affected information search, and found a correlation between short-term memory and the ratio of documents classified as irrelevant. MacFarlane et al. [27] reported that users with a high phonological memory capacity judged more documents as irrelevant compared to people with reduced capacity.

Berget and Sandnes [5] investigated the usability of an online search system with no query-building aids or tolerance for spelling errors. Dyslexia had a negative effect on search performance. It was suggested that search systems should compensate for misspelled queries to adequately accommodate users with dyslexia. This finding was supported in a second study by Berget and Sandnes [44] on query formulation in Google, a system with several query-building aids and a high tolerance level for errors. Users with dyslexia did not apply query-building aids significantly more than controls. However, the high tolerance for errors counteracted the negative effect of dyslexia.

3. WCAG 2.0

WCAG 2.0 has become the de facto standard for Web accessibility among developers, designers and legislators [45]. The first version was endorsed as a W3C recommendation in 1999 [46], but was replaced by WCAG 2.0 in 2008 [13], which is now defined as the international standard ISO/IEC 40500:2012 [47]. WCAG has been criticised, among others for placing too much responsibility on the end-users and require that they are aware of technologies that may best accommodate their needs [48].

Four principles constitute the basis for WCAG, namely that Web content should be perceivable, operable, understandable and robust. A total of 12 guidelines are formulated, one to four per principle. Each guideline is divided into success criteria, which are more specific, testable and platform independent requirements [13]. One of three conformance levels are defined for each criterion, from A to AAA, where AAA is the highest level. Guidelines accommodating people with dyslexia are found in all the principles. However, robustness is not discussed herein as it applies to assistive technology. Table 1 displays the main criteria that may particularly relates to dyslexia.

3.1 Navigation

Navigation is addressed in two guidelines, 2.4 and 3.2. Guideline 2.4 requires navigable content. The success criteria address how content is presented to support navigation, by allowing users to bypass blocks that are repeated on several Web pages (2.4.1) or by providing descriptive titles, headings and labels (2.4.2, 2.4.6 and 2.4.10). These measures may reduce the text that must be read, by either skipping repeated blocks, or by reading headings only, to locate the correct content. Other criteria address the need to provide information about the user's location within a Web site (2.4.8) and consistent navigation (3.2.3). Such measures may be especially important for users with dyslexia, who are reported to get confused when navigating Web pages [17,19]. Only criteria 2.4.1 (bypass blocks) and 2.4.2 (titles) are at the lowest conformity level.

3.2 Colours and Contrasts

Three success criteria (1.4.3, 1.4.6 and 1.4.8) address colours or contrasts, and are at level AA and AAA. Although visually impaired and colour blind users are specifically mentioned in the documentation [49], sufficient contrasts are also suggested to accommodate users with dyslexia [31].

3.3 Fonts and Graphics

WCAG does not give advice on the choice of fonts. However, two success criteria at level AA and AAA regard font size (1.1.4 and 1.4.8). Text should be possible to resize up to 200% without losing functionality, content or require horizontal scrolling. Most criteria addressing graphics emphasise that such content should not be the only source for information, but applied as decoration. Graphic size is barely mentioned in the supporting documentation for 1.1.4, with an emphasis on challenges related to images that may not rescale as well as text [50]. Criterion 1.4.8 mentions layout and discusses text width and line spacing, which may affect readability for users with dyslexia [35].

Table 1. Guidelines and success criteria relevant for users with dyslexia

Topic	Guideline	Success criteria	Description	Level
Navigation	2.4 Navigable	2.4.1	Bypass Blocks	A
		2.4.2	Page Titles	A
		2.4.6	Headings and Labels	AA
		2.4.8	Location	AAA
		2.4.10	Section Headings	AAA
	3.2 Predictable	3.2.3	Consistent Navigation	AA
Colours/contrasts	1.4 Distinguishable	1.4.3	Contrast (Minimum)	AA
		1.4.6	Contrasts (Enhanced)	AAA
		1.4.8	Visual Presentation	AAA
Font	1.4 Distinguishable	1.4.4	Resize Text	AA
		1.4.8	Visual Presentation	AAA
Layout	1.4. Distinguishable	1.4.8	Visual Presentation	AAA
Language	3.1 Readable	3.1.5	Reading Level	AAA
Timing	2.2 Enough Time	2.2.1	Timing Adjustable	A
		2.2.3	No Timing	AAA
Errors	3.3 Input Assistance	3.3.1	Error Identification	A
		3.3.3	Error Suggestion	AA
		3.3.5	Help	AAA
		3.3.6	Error Prevention	AAA

3.4 Language

Guideline 3.1 requires readable and understandable text. According to 3.1.5, a AAA criterion, the reading level should not exceed lower secondary education level [13]. This may be the criterion most clearly directed at users with dyslexia, who in addition to reading disabilities are specifically mentioned in the supporting document [51]. A short explanation of the difficulties experienced by people with dyslexia during reading are also provided.

3.5 Timing

Guideline 2.2 concerns timing. Several people with dyslexia have reduced writing or reading speed [15], which may cause difficulties when filling out forms or reading text if the user is not provided enough time. Two criteria address this issue. One criterion demands that users should be able to turn off or extend time limits (2.2.1, level A), while another states that there should be no timing (2.2.3, level AAA). The justification for this demand is, among others, to allow users with cognitive impairments enough time to read and understand text [52].

3.6 Errors

Guideline 3.3 deals with input assistance and errors where user input is required. One criterion (3.3.1) demands error identification, and suggests to provide users with information about required fields that has no input, or if input errors are detected. Other criteria concern error suggestions (3.3.3), context-sensitive help (3.3.5), error prevention and that data entered by the user is checked for errors (3.3.6). Users with reading disabilities are mentioned in the support document for the latter [53]. All the criteria are at a medium or high conformance level, except for 3.3.1.

4. Dyslexia-Specific Guidelines

In contrast to WCAG, guidelines directed at particular user groups exist, such as users with visual impairments [54], elderly people [55] or people with dyslexia [35, 56]. Friedman and Bryen [57] reviewed 20 guidelines directed at users with cognitive impairments, where three of these specifically addressed dyslexia. The most frequently appearing guideline (75% of the guidelines) regarded inclusion of graphic content in addition to text. Other commonly addressed topics were consistent design (60%) and clear text (60%). None of the 22 most frequently appearing guidelines addressed search systems, except for a general point regarding alerting users on possible errors.

Santana et al. [58] reviewed and summarised 41 dyslexia-specific guidelines. These guidelines covered navigation, colours, text presentation, writing, layout, images and charts, end user customization, mark-up, videos and audios. In these guidelines internal search and writing aids were included. Further, Santana et al. [58] includes a guideline stating that images and pictures should complement textual information. The justification for this guideline is that people with dyslexia consider images over words. However, the impact of adding graphic to text is disputed in the research literature [39].

The British Dyslexia Association [35] style guide consists of three parts: dyslexia friendly text, accessible formats and Website design. The first two address font types and sizes, and emphasise short and simple text. The third section accentuates clear navigation, and the use of graphics, images and pictures to break up text. A similar set of ten guidelines has been presented by Zarach [56]. These guidelines suggest using images alongside text, sans-serif fonts, includes a recommended font size, emphasise clear text, navigation and customisation of colours.

5. Discussion

WCAG seems to include many guidelines that may accommodate people with dyslexia (see Table 1). Several of the criteria are in accordance with difficulties reported by The Disability Rights Commission [19], such as navigation, colours and language. These issues are quite general, and may therefore also apply to people without dyslexia. Clear navigation, properly sized content and clear language will most likely positively affect the usability for all users. This is in accordance with previous findings, that measures which will improve accessibility for users with dyslexia can also enhance the usability for users without dyslexia [43].

Although several important topics for users with dyslexia are addressed in WCAG, a shortcoming is in the conformance levels. Only 23.5% of the success criteria in Table 1 are at the lowest level, while 29.4% are AA and 47.1% are AAA. Only certain criteria related to navigation, timing and errors are at the lowest conformance level. The criterion most clearly directed towards users with dyslexia, namely 3.1.5 regarding a clear and simple language, is classified as level AAA. Consequently, if jurisdictions and specifications of requirements do not apply the strictest conformance level, a high number of Web sites claiming to be in accordance with WCAG may be inaccessible for people with dyslexia. It may therefore be a need to rethink the classification of certain criteria, unless the highest level is always applied as formal requirements.

Existing guidelines directed at people with dyslexia focus on similar issues as WCAG, for instance colours, navigation and simple text. Fonts and images however, are not addressed in WCAG. Tolerance levels for misspelling are not included in any of the guidelines, although demands for correct spelling seem to impact people with dyslexia, particularly during search [5]. It may be argued that content guidelines are most clearly directed at the information made available online and not on user inputs. Consequently, misspellings made by users may be outside the scope of WCAG. However, the Web has become highly interactive, and most Web sites have some sections that require user input, such as forms for ordering products or tickets, systems for paying online or search on a Web page or in another type of search user interface.

Font typefaces and font characteristics have been discussed in the research literature. It has been reported that font types affect reading performance [32], that people with dyslexia prefer fonts without serifs [31] and that italics should be avoided [32]. Consequently, it may be useful to include a criterion regarding fonts, since WCAG now only seems to address font size. However, there may be conflicting needs with other user groups, so this must be discussed in a wider user diversity context.

The suggestion to include graphics in addition to text is not included in WCAG, but are mentioned in dyslexia-specific guidelines [35,56,58]. However, scholars do not agree upon the usefulness of dual-modality displays because an increased number of objects may cause a negative effect on cognitive load and performance measures [59].

It has been suggested that people with dyslexia struggle with information retrieval [4-5, 27]. Berget and Sandnes [5] found that search systems with no tolerance for errors are inaccessible to users with dyslexia. This finding is in accordance with Al-Wabil et al.'s findings [17]. Consequently, this is an important issue that should addressed.

WCAG should more explicitly incorporate search. Guideline 3.3 regards input assistance, but is mostly directed at filling out forms, and that feedback should be given when required fields are empty or if the inputted data does not seem correct. Nothing is mentioned about spelling errors in general, which are commonly made by users with dyslexia [15] and frequently occurs during query formulation [5]. Spelling errors in queries are reported to be common among users without dyslexia [60], which indicate that a guideline related to spell checking may also accommodate these users.

6. Conclusion

Dyslexia is rarely mentioned in WCAG, which may lead to neglect of the needs of users with reading disabilities [58]. W3C are aware of shortcomings regarding dyslexia in the introduction to the guidelines [13]. Consequently, more effort should be put into accommodating users with cognitive impairments. Much relevant research has been carried out regarding the needs of users with dyslexia. This platform of knowledge should constitute a suitable starting point for a discussion of the revision of WCAG.

WCAG should be revised to better accommodate people with dyslexia. The conformance levels should be altered to ensure that at least the criterion regarding clear language is at level A. It may also be useful to include criteria regarding font typefaces and text decoration. Search systems should be incorporated in WCAG more explicitly, including a guideline addressing the need for more assistance in user interfaces that require user input such as queries. Ideally, there should also be a demand for a certain tolerance level for misspellings in user input. Such criteria may be included in the existing guideline 3.3, and would not require a major revision of WCAG structure.

One may argue that by applying a combination of WCAG and a set of dyslexia-specific guidelines the accessibility increases. However, one of the purposes behind WCAG is to counteract the need to fulfil requirements in several guidelines. The overall aim is to develop one common standard which accommodates as many users as possible, and thus to achieve universal design. In this context, WCAG is a good starting point. Since dyslexia is such a widespread impairment, it seems reasonable to accommodate this user group better in future versions of WCAG.

References

[1] M. Dörk, C. Williamson, & S. Carpendale, Navigating tomorrow's Web: From searching and browsing to visual exploration, *ACM Transactions on the Web* **6** (2012), article no. 13.

[2] T. Berners-Lee & R. Cailliau, World-Wide Web, In *Computing in High Energy Physics 92*, Annecy, France, 23-27 September 1992, Retrieved from http://www.freehep.org/chep92www.pdf.

[3] T.Gossen, J. Hempel & A. Nürnberger, Find it if you can: Usability case study of search engines for young users, *Personal and Ubiquitous Computing* **17** (2013), 1593-1603.

[4] A. MacFarlane, A. Al-Wabil, C.R. Marshall, A. Albrair, S.A. Jones & P. Zaphiris, The effect of dyslexia on information retrieval: A pilot study, *Journal of Documentation*, **66** (2010), 307-326.

[5] G. Berget & F.E. Sandnes, Searching databases without query-building-aids: Implications for dyslexic users, *Information Research* **20** (2015), article 689.

[6] Center for Universal Design, *About UD,* 2008, Retrieved 13.05.2016 from https://www.ncsu.edu/ncsu/design/cud/about_ud/about_ud.htm

[7] UN, *Convention on the rights of persons with disabilities*, 2006, Retrieved 13.05.2016 from http://www.un.org/disabilities/convention/conventionfull.shtml.

[8] European Parliament, *eEurope 2002: Accessibility of Public Web Sites and their Content: European Parliament resolution on the Commission communication eEurope 2002: Accessibility of Public Web Sites and their Content*, 2002, Retrieved 13.05.2016 from http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+TA+P5-TA-2002-0325+0+DOC+PDF+V0//EN

[9] Kommunal- og Moderniseringsdepartementet, *Forskrift om universell utforming av informasjons- og kommunikasjonsteknologiske løsninger* [Regulation on universal design of information and communation technological solutions], 2013, Retrieved 13.05.2016 from https://lovdata.no/dokument/SF/forskrift/2013-06-21-732?q=Forskrift%20om%20universell%20utforming%20av*

[10] Treasury Board of Canada, *Standard on Web Accessibility*, 2013, Retrieved 13.05.2016 from http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?section=text&id=23601

[11] Australian Human Rights Commission, *World Wide Web Access: Disability Discrimination Act Advisory Notes ver. 4.0*, 2014, Retrieved 13.05.2016 from http://humanrights.gov.au/world-wide-web-access-disability-discrimination-act-advisory-notes-ver-40-2010#transition

[12] R. Huffaker, Enforcing eAccessibility: Is the current legal framework adequate?, *International Review of Law, Computers & Technology* **29** (2015), 207-225.

[13] W3C, *Web content accessibility guidelines (WCAG) 2.0*, 2008, Retrieved 15.05.2016 from http://www.w3.org/TR/WCAG20/.
[14] A.F. Newell, & P. Gregor, "User Sensitive Inclusive Design": In search of a new paradigm. In J. Thomas (Ed.), *CUU '00 Proceedings on the 2000 conference on Universal Usability* (pp. 39-44). New York: ACM, 2000.
[15] M.J. Snowling, *Dyslexia*, Blackwell, Malden, 2000.
[16] M. Ahissar, Dyslexia and the anchoring-deficit hypothesis, *Trends in Cognitive Sciences* **11** (2007), 458-465.
[17] A. Al-Wabil, P. Zaphiris & S. Wilson, Web navigation for individuals with dyslexia: An exploratory study. In C. Stephanidis (Ed.), *Universal Access in Human Computer Interaction: Coping with Diversity: Lecture Notes in Computer Science, 4th International Conference on Universal Access in Human-Computer Interaction, UAHCI 2007, Held as Part of HCI International 2007, Beijing, China, July 22-27, 2007, Proceedings, Part I* (pp. 593–602), Springer, Berlin, 2007.
[18] D. Rømen, & D. Svanæs, Validating WCAG versions 1.0 and 2.0 through usability testing with disabled users, *Universal Access in the Information Society* **11** (2012), 375-385.
[19] Disability Rights Commission, The Web: *Access and inclusion for disabled people: A formal investigation conducted by the Disability Rights Commission*, TSO, London, 2004.
[20] L. Habib, G. Berget, F.E. Sandnes, N. Sanderson, P. Kahn, S. Fagernes & A. Olcay, Dyslexic students in higher education and virtual learning environments: An exploratory study, *Journal of Computer Assisted Learning* **28** (2012), 574-584.
[21] M. van der Schoot, R. Licht, T.M. Horsley & J.A. Sergeant, Inhibitory deficits in reading disability depend on subtype: Guessers not spellers. *Child Neuropsychology* **6** (2000), 297-312.
[22] S.E. Shaywitz & B.A. Shaywitz, Dyslexia (specific reading disability). *Biological Psychiatry* **57** (2005), 1301–1309.
[23] J.M. Caroll & J.E. Iles, An assessment of anxiety levels in dyslexic students in higher education, *British Journal of Educational Psychology* **76** (2006), 651-662.
[24] T. Mortimore & W.R. Crozier, Dyslexia and difficulties with study skills in higher education, *Studies in Higher Education* **31** (2006), 235-251.
[25] J.H. Smith-Spark & J.E. Fisk, Working memory functioning in developmental dyslexia, *Memory* **15** (2007), 34-56.
[26] R. Nicolson & A. Fawcett, Automaticity: A new framework for dyslexia research?, *Cognition* **35** (1990), 159–182.
[27] A. MacFarlane, A. Albrair, C.R. Marshall, & G. Buchanan, Phonological working memory impacts on information searching: An investigation of dyslexia. In J. Kamps, W. Kjraaj, N. Fuhr (Eds.) *Proceedings of IIiX 2012 the Fourth Information Interaction in Context Symposium, Nijmegen, the Netherlands, August 21-24, 2012* (pp 27-34), ACM, New York, 2012.
[28] S. Kurniawan & G.V. Conroy, In S. Kurniawan & P. Zaphiris (Eds.), *Advances in Universal Web Design and Evaluation* (pp. 271-292), Idea Group, Hershey, 2007.
[29] L. Rello, G. Kanvinde & R. Baeza-Yates, Layout guidelines for web text and a web service to improve accessibility for dyslexics. In M. Vigo, J. Abascal, R. Lopes & P. Salomoni (Eds.), *W4A'12 Proceedings of the International Cross-Disciplinary Conference on Web Accessibility* (article no. 36), ACM, New York, 2012.
[30] J. Nielsen, Miniatures versus icons as visual cache for videotext browsing, *Behaviour & Information Technology* **9** (1990), 441-449.
[31] L. Evett & D. Brown, Text formats and Web design for visually impaired and dyslexic readers: Clear text for all, *Interacting With Computers*, **17** (2005), 453-472.
[32] L. Rello & R. Baeza-Yates, Good fonts for dyslexia. In *ASSESTS '13 Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (article no. 14), ACM, New York, 2013.
[33] L. Rello, M. Pielot, M.-C. Marcos & R. Carlini, Size matters (spacing not): 18 points for a dyslexic-friendly Wikipedia, In G. Brajnik & P. Salomoni (Eds.), *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility, May 13–15, 2013, Rio de Janeiro, Brazil (W4A'13)* (article no. 17), ACM, New York, 2013.
[34] M. Zorzi, C. Barbiero, A. Facoetti, I. Lonciari, M. Carrozzi, M. Montico, … J.C. Ziegler, Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences of the United States of America* **109** (2012), 11455-11459.
[35] British Dyslexia Association, *Dyslexia Style Guide*, n.d., Retrieved 13.05.2016, from http://www.bdadyslexia.org.uk/common/ckeditor/filemanager/userfiles/About_Us/policies/Dyslexia_Style_Guide.pdf
[36] S. Murphy, Accessibility of graphics in technical documentation for cognitive and visually impaired. In S. Tilley & R. M. Newman (Eds.), *SIGDOC'05 Proceedings of the 23rd annual international*

conference on Design and communication: Documenting & designing for pervasive information (pp. 12-17), ACM, New York, 2005.
[37] S. Kim, L.J. Lombardino, W. Cowles & L.J. Altman, Investigating graph comprehension in students with dyslexia: An eye tracking study, *Research in Developmental Disabilities* **35** (2014), 1609-1622.
[38] N.A. Beacham, & J.L. Alty, An investigation into the effects that digital media can have on learning outcomes of individuals who have dyslexia, *Computers & Education* **47** (2006), 74-93.
[39] E.W. Brante, M.H. Olander & M. Nyström, Exploring the impact of contrasting cases in text and picture processing, *Journal of Visual Literacy* **23** (2013), 15-38.
[40] G. Berget, F. Mulvey & F.E. Sandnes, Is visual content in textual search interfaces beneficial to dyslexic users?, *International Journal of Human-Computer Studies* **92-93** (2016), 17-29.
[41] L. Rello, R. Baeza-Yates, L. Dempere-Marco & H. Saggion, Frequent words improve readability and short words improve understandability for people with dyslexia, In P. Kotzé et al. (Eds.), *Human-Computer Interaction – INTERACT 2013, Volume 8120 of the Series Lecture Notes in Computer Science* (pp. 203-219), Springer, Berlin, 2013.
[42] C. Boldyreff, E. Burd, J. Donkin & S. Marshall, The case for the use of plain English to increase Web accessibility, In S. Tilley (Ed.), *Proceedings of the 3rd International Workshop on Web Site Evolution* (pp. 42-48), IEEE, Los Alamitos, 2001.
[43] J.E. McCarthy & S.J. Swierenga, What we know about dyslexia and Web accessibility: A research review, *Universal Access in the Information Society* ***9*** (2010), 147-152.
[44] G. Berget & F.E. Sandnes, Do autocomplete functions reduce the effect of dyslexia on information searching behavior? The case of Google, *Journal of the American Society for Information Science & Technology* (n.d.), accepted 2015.
[45] L. Moreno, P. Martínez & B. Ruiz-Mezcua, Disability standards for multimedia on the Web, *MultiMedia IEEE* **15** (2008), 52-54.
[46] W3C, *Web Content Accessibility Guidelines 1.0*, 1999, Retrieved 13.05.2016 from http://www.w3.org/TR/WAI-WEBCONTENT/.
[47] ISO, *ISO/IEC 40500:2012: Information technology – W3C Web Content Accessibility Guidelines (WCAG) 2.0*, ISO, Geneva, 2012.
[48] D. Sloan, A. Heath, F. Hamilton, B. Kelly, H. Petrie & L. Phipps, Contextual Web accessibility: Maximizing the benefit of accessibility guidelines. In S. Harper & Y. Yesilada (Eds.), *W4A '07 Proceedings of the 2007 international cross-disciplinary conference on Web accessibility*, ACM, New York, 2007.
[49] W3C, *Contrast (Enhanced): Understanding SC 1.4.6*, 2016a, Retrieved 18.05.2016 from http://www.w3.org/TR/UNDERSTANDING-WCAG20/visual-audio-contrast7.html.
[50] W3C, *Resize text: Understanding SC 1.4.4*, 2016b, Retrieved 18.05.2016 from http://www.w3.org/TR/UNDERSTANDING-WCAG20/visual-audio-contrast-scale.html.
[51] W3C, *Reading level: Understanding SC 3.1.5*, 2016c, Retrieved 18.05.2016 from http://www.w3.org/TR/UNDERSTANDING-WCAG20/meaning-supplements.html.
[52] W3C, *No timing: Understanding SC 2.2.3*, 2016d, Retrieved 18.05.2016 from http://www.w3.org/TR/UNDERSTANDING-WCAG20/time-limits-no-exceptions.html.
[53] W3C, *Error identification: Understanding SC 3.3.1*, 2016e, Retrieved 18.05.2016 from http://www.w3.org/TR/UNDERSTANDING-WCAG20/minimize-error-identified.html.
[54] S. Leuthold, J.A. Bargas-Avila & K. Opwis, Beyond Web content accessibility guidelines: Design of enhanced text user interfaces for blind internet users, *International Journal of Human-Computer Studies* **66** (2008), 257-270.
[55] P. Zaphiris, S. Kurniawan & M. Ghiawadwala, A systematic approach to the development of research-based Web design guidelines for older people, *Universal Access in the Information Society* **6** (2007), 59-75.
[56] V. Zarach, *Ten guidelines for improving accessibility for people with dyslexia*, 2002, Retrieved 14.05.2016 from http://wiki.cetis.ac.uk/Ten_Guidelines_for_Improving_Accessibility_for_People_with_Dyslexia
[57] M.G. Friedman & D.N. Bryen, Web accessibility design recommendations for people with cognitive disabilities, *Technology and Disability* **19** (2007), 205-212.
[58] V.F. de Santana, R. de Oliveira, L.D.A. Almeida & M.C.C. Baranauskas, Web accessibility and people with dyslexia: A survey on techniques and guidelines, In M. Vigo et al. (Eds.), *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility (W4A'12) Lyon, France, 16–17 April* (article no.35), ACM, New York, 2012.
[59] E. Moores, R. Cassim & J.B. Talcott, Adults with dyslexia exhibit large effects of crowding, increased dependence on cues, and detrimental effects of distractors in visual search tasks, *Neuropsychologica* **49** (2011), 3881-3890.

[60] S. Cucerzan & E. Brill, Spelling correction as an iterative process that exploits the knowledge of Web users. In D. Ling & D. Wu (Eds.), *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, EMNLP 2004* (pp. 293-300), ACL, Barcelona 2004.

Appendix G

Title: Ethics approval letter from the Norwegian Social Science Data Services (NSD)

Norsk samfunnsvitenskapelig datatjeneste AS
NORWEGIAN SOCIAL SCIENCE DATA SERVICES

Harald Hårfagres gate 29
N-5007 Bergen
Norway
Tel: +47-55 58 21 17
Fax: +47-55 58 96 50
nsd@nsd.uib.no
www.nsd.uib.no
Org.nr. 985 321 884

Gerd Berget
Institutt for informasjonsteknologi
Høgskolen i Oslo og Akershus
Postboks 4 St. Olavs plass
0130 OSLO

Vår dato: 06.02.2012 Vår ref: 29348 / 3 / LT Deres dato: Deres ref:

TILRÅDING AV BEHANDLING AV PERSONOPPLYSNINGER

Vi viser til melding om behandling av personopplysninger, mottatt 11.01.2012. Meldingen gjelder prosjektet:

29348	*Studenter med dysleksi i interaksjon med søkesystemer - på vei mot universelt utformede grensesnitt?*
Behandlingsansvarlig	*Høgskolen i Oslo og Akershus, ved institusjonens øverste leder*
Daglig ansvarlig	*Gerd Berget*

Personvernombudet har vurdert prosjektet, og finner at behandlingen av personopplysninger vil være regulert av § 7-27 i personopplysningsforskriften. Personvernombudet tilrår at prosjektet gjennomføres.

Personvernombudets tilråding forutsetter at prosjektet gjennomføres i tråd med opplysningene gitt i meldeskjemaet, korrespondanse med ombudet, eventuelle kommentarer samt personopplysningsloven/-helseregisterloven med forskrifter. Behandlingen av personopplysninger kan settes i gang.

Det gjøres oppmerksom på at det skal gis ny melding dersom behandlingen endres i forhold til de opplysninger som ligger til grunn for personvernombudets vurdering. Endringsmeldinger gis via et eget skjema, http://www.nsd.uib.no/personvern/forsk_stud/skjema.html. Det skal også gis melding etter tre år dersom prosjektet fortsatt pågår. Meldinger skal skje skriftlig til ombudet.

Personvernombudet har lagt ut opplysninger om prosjektet i en offentlig database, http://www.nsd.uib.no/personvern/prosjektoversikt.jsp.

Personvernombudet vil ved prosjektets avslutning, 01.01.2016, rette en henvendelse angående status for behandlingen av personopplysninger.

Vennlig hilsen

Vigdis Namtvedt Kvalheim

Lis Tenold

Kontaktperson: Lis Tenold tlf: 55 58 33 77
Vedlegg: Prosjektvurdering

Personvernombudet for forskning

Prosjektvurdering - Kommentar

Prosjektnr: 29348

Det gis skriftlig informasjon og innhentes skriftlig samtykke. Personvernombudet finner skrivet godt utformet.

Personvernombudet finner at det samles inn og registreres sensitive personopplysninger om helseforhold, jf. personopplysningsloven § 2 nr. 8 bokstav c.

Innsamlede opplysninger anonymiseres og lydbåndopptak makuleres ved prosjektslutt, senest 01.001.2016. Med anonymisering innebærer at navnelister slettes/makuleres, og ev. kategorisere eller slette indirekte personidentifiserbare opplysninger. Ved publisering vil ingen enkeltpersoner kunne gjenkjennes.