DIGITAL TEXT PRESENTATION AND NAVIGATION TO SUPPORT PEOPLE WITH DYSLEXIA

Ili Farhana Binti Md Mahtar

Doctor of Philosophy

UNIVERSITY OF YORK

COMPUTER SCIENCE

DECEMBER 2018

ABSTRACT

Dyslexia is a reading disability that is characterised by difficulties of reading, decoding and spelling. In order make online materials accessible for people with dyslexia, developers should make on-screen presentation of text adaptable. There is very little research that has empirically tested which text presentations and web navigation are helpful or acceptable to people with dyslexia. Therefore two studies are conducted on the aspect of text presentation, Study 1 focused on the effects of typefaces and font size while Study 2 focused on the effects of line spacing and line length. Study 3 focused on the effects on menu organisation and visibility on web navigation. All three studies compared English native speaking adults with and without dyslexia on their eye gaze behaviour, performance, preferences and opinions. For the text presentation studies, the dyslexic participants were grouped into more specific categories, mild and moderately dyslexic, based on the results of a well-established checklist for identifying dyslexia. Eye gaze tracking was measured in all studies. Findings from the studies on text presentation show that all participants had fewer fixations with small font size, shortest fixation durations with dyslexia-optimised typefaces, and fewer and shorter fixations with longer line length. Participants preferred sans serif typeface and wider line spacing. There were different levels of comfort with dyslexia-optimised typefaces for non-dyslexic and dyslexic participants. Findings from the study on web navigation show that fragmented menus with visible sub-menus had fewer fixations, while dynamic sub-menus had fewer revisits and fewer fixations. However unified menus were rated as easier to use, to remember and to learn. Participants with dyslexia show poor performance in both text presentation and web navigation studies. Key contribution of this programme of research is to the methodology of studies to investigate text presentation on screen and web navigation effects for people with dyslexia.

LIST OF CONTENTS

ABSTRACT **1**

LIST OF CONTENTS **3**

LIST OF FIGURES **6**

LIST OF TABLES **10**

ACKNOWLEDGEMENTS **13**

DECLARATION **14**

CHAPTER 1: INTRODUCTION TO RESEARCH AND DYSLEXIA **15**

1.1 INTRODUCTION 15
1.2 MOTIVATION FOR THIS RESEARCH 16
1.3 WORD RECOGNITION PROCESS IN READING 16
1.4 DYSLEXIA: THEORIES, DEFINITION AND CAUSE 18
1.5 CHARACTERISTICS OF DYSLEXIA 21
1.6 THE PREVALENCE OF DYSLEXIA 23
1.7 THE CONTROVERSY AND MISCONCEPTION SURROUNDING DYSLEXIA 24
1.8 AIMS AND SCOPE OF RESEARCH 25
1.9 STRUCTURE OF THE THESIS 28

CHAPTER 2: LITERATURE REVIEW **30**

2.1 INTRODUCTION 30
2.2 TEXT PRESENTATION AND TYPEFACES DESIGNED FOR PEOPLE WITH DYSLEXIA 30
2.2.1 Characteristics of Typefaces 30
2.2.2 Dyslexia-Optimised Typefaces 32
2.2.3 Criticism on Dyslexia-Optimised Typefaces 35
2.3 IMPACT OF TEXT PRESENTATION ON READING BEHAVIOUR OF PEOPLE WITH DYSLEXIA 38
2.3.1 Research on the Effects of Typefaces 38
2.3.2 Research on the Effects of Font Sizes 45
2.3.3 Research on the Effects of Line Lengths 52
2.3.4 Research on the Effects of Line Spacings 54
2.3.5 Research on the Effects of Colours 59
2.4 THE WAY OF PEOPLE WITH DYSLEXIA NAVIGATING WEB SITES 63
2.5 GUIDELINES AND RECOMMENDATIONS PUBLISHED FOR PEOPLE WITH DYSLEXIA 68
2.6 CURRENT ASSISTIVE TECHNOLOGIES CREATED FOR PEOPLE WITH DYSLEXIA 78
2.7 DIFFICULTIES ENCOUNTERED BY PEOPLE WITH DYSLEXIA WHEN USING WEB SITES 79
2.7.1 Unclear and Confusing Web Page Layout 79
2.7.2 Unclear and Confusing Navigation Mechanisms 79
2.7.3 Inappropriate Use of Colours 80
2.7.4 Small Graphics and Text 81
2.7.5 Complicated Language 82

2.8 Conclusion 82

CHAPTER 3: STUDY 1 – EFFECTS OF FONT SIZE AND TYPEFACE ON DETAILED READING 90

3.1 Introduction 90
3.2 Preparatory Study – Assessment of Questions Difficulty 91
3.2.1 Method 92
3.2.2 Result and Discussion 94
3.3 Study 1 98
3.3.1 Method 100
3.3.2 Data Preparation 110
3.3.3 Results 112
3.3.4 Discussion and Conclusion 130

CHAPTER 4: STUDY 2 – EFFECTS OF LINE SPACING AND LINE LENGTH ON DETAILED READING 137

4.1 Introduction 137
4.2 Preparatory Study – Assessment of Question Difficulty 138
4.2.1 Method 138
4.2.2 Results and Discussion 140
4.3 Study 2 142
4.3.1 Method 142
4.3.2 Data Preparation 150
4.3.3 Results 151
4.3.4 Discussion and Conclusion 161

CHAPTER 5: STUDY 3 – EFFECTS OF MENU ORGANISATION AND VISIBILITY ON WEB NAVIGATION 165

5.1 Introduction 165
5.2 Preparatory Study – Assessment of Question Difficulty 166
5.2.1 Method 166
5.2.2 Results and Discussion 169
5.3 Study 3 170
5.3.1 Method 171
5.3.2 Data Preparation 181
5.3.3 Results 183
5.3.4 Discussion and Conclusion 204

CHAPTER 6: OVERALL DISCUSSION AND CONCLUSIONS 212

6.1 Introduction 212
6.2 Overview of the Programme of Research 212
6.3 Contributions of the Programme of Research 219
6.3.1 Improvements to the Methodology of Studies in Investigating Digital Text Presentation and Navigation 219
6.3.2 Demonstration of the Variability of Results 222

6.3.3 *Contributions to the Development and Updating of Web Accessibility Guidelines for People with Dyslexia* ... 226
6.3.4 *Other Contributions* ... 227
6.4 Discussion of Research Questions ... 227
6.5 Limitations and Future Work ... 228
6.5.1 *Further Investigations on Longer Texts and Longer Vertical Menus* ... 229
6.5.2 *Further Investigations on Mouse Movement in Parallel with Eye Gaze* ... 230
6.5.3 *Analysing the Scan Paths in Web Navigation* ... 230
6.5.4 *Extending the Research to Different Categories of Dyslexic Participants* ... 231
6.6 Conclusions ... 231

APPENDIX A – CONVERSION OF LOGMAR PRINT SIZES ... 233

APPENDIX B – SAMPLE OF TEXT, QUIZ AND DIFFICULTY ASSESSMENT FOR STUDY 1 ... 234

APPENDIX C – ADULT READING QUESTIONNAIRE (ARQ) ... 250

APPENDIX D – INFORMED CONSENT FORM ... 252

APPENDIX E – PRE-STUDY QUESTIONNAIRE ... 254

APPENDIX F – POST-STUDY QUESTIONNAIRE (NON-DYSLEXIC) FOR STUDY 1 ... 255

APPENDIX G – POST-STUDY QUESTIONNAIRE (DYSLEXIC) FOR STUDY 1 ... 257

APPENDIX H – SEQUENCE OF TEXTS FOR PRACTICE AND EXPERIMENT TASKS IN STUDY 1 ... 260

APPENDIX I – DESCRIPTIVE STATISTICS FOR RATINGS OF EASE OF READING IN STUDY 1 ... 261

APPENDIX J – SEQUENCE OF TEXTS CREATED USING REPLICATED LATIN SQUARE DESIGN (RLSD) IN STUDY 2 ... 263

APPENDIX K – SAMPLE OF TEXT, QUIZ AND DIFFICULTY ASSESSMENT FOR STUDY 2 ... 265

APPENDIX L – POST STUDY QUESTIONNAIRE FOR STUDY 2 ... 271

APPENDIX M – QUESTIONS AND TASKS DIFFICULTY ASSESSMENT FOR STUDY 3 ... 273

APPENDIX N – FIRST POST-STUDY QUESTIONNAIRE FOR STUDY 3 ... 276

APPENDIX O – SECOND POST-STUDY QUESTIONNAIRE FOR STUDY 3 ... 278

APPENDIX P – SEQUENCE OF QUESTIONS FOR EXPERIMENT TASKS IN STUDY 3 ... 280

REFERENCES ... 281

LIST OF FIGURES

Figure 2.1 Example of text with Proportional (Helvetica, 14pt) and Mono-spaced (Courier, 14pt) typefaces ... 31

Figure 2.2 Example of text Roman, Italic and Oblique shapes (all Garamond typeface) ... 31

Figure 2.3 Example of Serif and Sans Serif categories of typefaces ... 31

Figure 2.4 Samples of text using Arial (14pt), OpenDyslexic (14pt) and Dyslexie (14pt) typefaces ... 32

Figure 2.5 Sample of text in Comic Sans MS (14pt) and Lexie Readable typefaces (14pt) ... 32

Figure 2.6 Unique letter shapes and heavier bottom of OpenDyslexic (Source: DyslexieFont.com, n.d.) ... 35

Figure 2.7 Effect of colour-typeface combinations on reading speed for participants with dyslexia (Source: Pijpker, 2013) ... 40

Figure 2.8 Effect of colour-typeface combinations on reading accuracy for participants with dyslexia (Source: Pijpker, 2013) ... 41

Figure 2.9 LogMAR chart (Source: Precision Vision, n.d.) ... 46

Figure 2.10 Example of reading speed for two children with and without dyslexia (Source: O'Brien et al., 2005) ... 47

Figure 2.11 Fixation duration by font size (Source: Rello, Pielot, et al., 2013) ... 48

Figure 2.12 Comprehension score by font size (Source: Rello, Pielot, et al., 2013) ... 49

Figure 2.13 User rating for readability of font sizes (Source: Rello, Pielot, et al., 2013) ... 49

Figure 2.14 User rating for comprehensibility of font sizes. (Source: Rello, Pielot, et al., 2013) ... 49

Figure 2.15 Reading performance and preferences of font size for participants with dyslexia (Group D) and without dyslexia (Group C) (Source: Rello & Baeza-Yates, 2015) ... 50

Figure 2.16 Text with default setting for control group (Source: Schoonewelle, 2013) ... 51

Figure 2.17 Text with altered setting for experiment group (Source: Schoonewelle, 2013) ... 51

Figure 2.18 Reading with iPod was more efficient compared to iPad (Source: Schneps, Thomson, Sonnert, et al., 2013) ... 53

Figure 2.19 Reading performance and preferences of line length for participants with dyslexia (Group D) and without dyslexia (Group C) (Source: Rello & Baeza-Yates, 2015) ... 54

Figure 2.20 Fixation duration by line spacing (Source: Rello, Pielot, et al., 2013) ... 56

Figure 2.21 Comprehension score by line spacing (Source: Rello, Pielot, et al., 2013) ... 56

Figure 2.22 Interaction between font size and line spacing (Source: Rello, Pielot, et al., 2013) ... 56

Figure 2.23 Reading performance and preferences of line spacing for participants with dyslexia (Group D) and without dyslexia (Group C) (Source: Rello & Baeza-Yates, 2015)....................57

Figure 2.24 Sample of text read by participants. (A) Normal condition, (B) Spaced condition (Source: Zorzi et al., 2012)....................58

Figure 2.25 Combinations of colour pairs with colour and brightness differences (Source: Rello & Baeza-Yates, 2012)....................59

Figure 2.26 User performance and user preferences of colour pairs (Source: Rello & Baeza-Yates, 2012)....................60

Figure 2.27 HEX code and luminosity contrast ratio for all colour pairs (Source: Rello & Bigham, 2017)....................61

Figure 2.28 Reading performance and preferences of text/background colours for participants with dyslexia (Group D) and without dyslexia (Group C) (Source: Rello & Baeza-Yates, 2015)....................63

Figure 2.29 Heat map showing region of interest for all participants (Source: Al-Wabil et al., 2008)....................66

Figure 2.30 Gaze plot of fixation before selecting target link for one participant (Source: Al-Wabil et al., 2008)....................66

Figure 2.31 Combination of all values in Rello, Kanvinde, et al.'s (2012b) guidelines....................72

Figure 2.32 Examples of some parameters including its conditions (Source: Rello & Baeza-Yates, 2015)....................73

Figure 3.1 Sample multiple choice question and answers....................93

Figure 3.2 Different size between absolute and relative unit (Source: Schaeffer, 2008)....................99

Figure 3.3 Different height for Times New Roman and Georgia....................99

Figure 3.4 Dual monitor setup for Study 1....................102

Figure 3.5 Sample of text used in Study 1....................108

Figure 3.6 Sample of comprehension questions used in Study 1....................108

Figure 3.7 Sample of accompanying web page used in Study 1....................109

Figure 3.8 Post-hoc of Group on number of fixations....................114

Figure 3.9 Interaction of Group and Font Size on number of fixations....................114

Figure 3.10 Interaction of Typeface and Font Size on number of fixations....................115

Figure 3.11 Interaction of Typeface on average fixation duration....................116

Figure 3.12 Interaction of Group and Typeface on average fixation duration....................117

Figure 3.13 Mean of Text Reading Time for all groups....................119

Figure 3.14 Ease of reading ratings for different typefaces and font sizes....................120

Figure 3.15 Ease of reading ratings of non-dyslexic participants for sans serif, serif and dyslexia-optimised typefaces 121

Figure 3.16 Ease of reading ratings of mild dyslexic participants for sans serif, serif and dyslexia-optimised typefaces 122

Figure 3.17 Distribution of ratings whether the dyslexia-optimised typefaces were easier to read than standard typefaces 125

Figure 3.18 Distribution of ratings of interest in using dyslexia-optimised typefaces 126

Figure 3.19 Distribution of comfort ratings by non-dyslexic participants on dyslexia-optimised typefaces 128

Figure 3.20 Distribution of comfort ratings of mild-dyslexic participants on dyslexia-optimised typefaces 128

Figure 3.21 Distribution of comfort ratings of moderate dyslexic participants on dyslexia-optimised typefaces 129

Figure 4.1 Sample of multiple choice question and answers 139

Figure 4.2 Sample of Experimental Text 1 with version A (line spacing of 1.5 and line length of 60 – 70cpl) used in Study 2 147

Figure 4.3 Sample of Experimental Text 1 with version B (line spacing of 2.0 and line length of 80 – 90cpl) used in Study 2 147

Figure 4.4 Sample of Experimental Text 1 with version C (line spacing of 1.5 and line length of 80 – 90cpl) used in Study 2 148

Figure 4.5 Sample of Experimental Text 1 with version D (line spacing of 2.0 and line length of 60 – 70cpl) used in Study 2 148

Figure 4.6 Sample of companion web page used in Study 2 149

Figure 4.7 Mean average fixation duration for the three participant groups 155

Figure 4.8 Mean text reading time for the three participant groups 156

Figure 4.9 Mean total comprehension score for the three participant groups 157

Figure 4.10 Means and standard deviations for ease of reading ratings for different line spacing and line length for the three participant groups 158

Figure 5.1 Sample of the web site in Preparatory Study 3 168

Figure 5.2 Hierarchical structure of menus and web pages in the web site 168

Figure 5.3 Sample of multiple choice questions with its alternatives 168

Figure 5.4 Sample of practice web site used in Study 3 175

Figure 5.5 Sample of experiment web site in Study 3 with Unified-Visible menus 176

Figure 5.6 Sample of experiment web site in Study 3 with Unified-Dynamic menus 176

Figure 5.7 Sample of experiment web site in Study 3 with Fragmented-Visible menus 177

Figure 5.8 Sample of experiment web site in Study 3 with Fragmented-Dynamic menus 177

Figure 5.9 Screenshot of question before navigating the web site in Study 3 178

Figure 5.10 Screenshot of question after navigating the web site in Study 3 179

Figure 5.11 Position of cue card in Study 3 179

Figure 5.12 Mean number of revisits for the two participant groups 190

Figure 5.13 Mean of number of fixations for all combinations of main menus and sub-menus 193

Figure 5.14 Possible revisit pattern in the Unified-Visible menu condition 207

Figure 5.15 Possible revisit pattern in the Unified-Dynamic menu condition 208

LIST OF TABLES

Table 2.1 Features of Dyslexie font (Source: DyslexieFont.com, n.d.) ... 33

Table 2.2 Sizes of Dyslexie, Times New Roman and Arial in 12-point size (Source: Kuster et al., 2017) ... 35

Table 2.3 Details of all conditions in a study with poor readers (Source: Marinus et al., 2016) ... 36

Table 2.4 Typefaces and font sizes used in all three conditions (Source: Kuster et al., 2017) ... 42

Table 2.5 Means (Standard deviations) by group and reading grade equivalent for children with and without dyslexia (Source: O'Brien et al., 2005) ... 47

Table 2.6 Task completion time on three conditions for dyslexia and control group (Source: Al-Wabil et al., 2008) ... 65

Table 2.7 Results on eye movement metrics between dyslexia and control group (Source: Al-Wabil et al., 2008) ... 66

Table 2.8 Results of Levenshtein distance for scan paths between participants with and without dyslexia (Source: Al-Wabil et al., 2008) ... 67

Table 2.9 Mapping between recommendation groups and relevance of stakeholder's roles (Source: de Santana et al., 2012) ... 69

Table 2.10 Results for each parameters value (Source: Rello et al., 2012b) ... 70

Table 2.11 Guidelines for the Web (Source: Rello et al., 2012b) ... 71

Table 2.12 Results and recommendations for more readable text (Source: Rello & Baeza-Yates, 2015) ... 74

Table 2.13 Summary of text presentation and colour pair from various guidelines ... 76

Table 2.14 Compilation of categories covered in guidelines ... 77

Table 2.15 Systematic literature review matrix ... 84

Table 3.1 Results for first round of preparatory study ... 95

Table 3.2 Results for second round of preparatory study ... 98

Table 3.3 Results of readability scores for all sets of texts used in Study 1 ... 105

Table 3.4 List of all font combinations in practice and experimental texts used for Study 1 ... 105

Table 3.5 Descriptive statistics of number of fixations for all participant groups ... 113

Table 3.6 Summary of three-way mixed ANOVA on number of fixations ... 113

Table 3.7 Descriptive statistics of average fixation duration for all participant groups ... 116

Table 3.8 Summary of three-way mixed ANOVA on average fixation duration ... 116

Table 3.9 Descriptive statistics of text reading time for all participant groups ... 118

Table 3.10 Summary of three-way mixed ANOVA on text reading time ... 118

Table 3.11 Distribution of preferences for typeface by the three participant groups ... 122

Table 3.12 Distribution of preferences for font size by the three participant groups ... 123

Table 3.13 Numbers of participants who had previously encountered on dyslexia optimised typefaces ... 123

Table 3.14 Descriptive statistics of dyslexic participants' opinions on dyslexia-optimised typefaces ... 124

Table 3.15 Descriptive statistics for dyslexics' comfortability on dyslexia-optimised typefaces ... 127

Table 4.1 Results for all texts in the preparatory study ... 141

Table 4.2 Results of readability scores for experimental texts used in Study 2 ... 146

Table 4.3 Descriptive statistics of number of fixations for all participant groups ... 153

Table 4.4 Summary of three-way mixed ANOVA on number of fixations ... 153

Table 4.5 Descriptive statistics for average fixation duration (ms) for all participant groups ... 154

Table 4.6 Three-way mixed ANOVA with repeated measures on average fixation duration ... 154

Table 4.7 Descriptive statistics for reading time (ms) for all participant groups ... 155

Table 4.8 Summary of three-way mixed ANOVA on reading time ... 156

Table 4.9 Ease of reading ratings for combinations of line spacing and line length ... 158

Table 4.10 Summary of three-way mixed ANOVA on ease of reading ... 159

Table 4.11 Distribution of preferences for line spacing by the three participant groups ... 160

Table 4.12 Distribution of preferences for line length by the three participant groups ... 160

Table 5.1 Result for all questions in the Preparatory Study ... 170

Table 5.2 Descriptive statistics of entry time (ms) for both participant groups in all AOIs ... 184

Table 5.3 Descriptive statistics of dwell time (ms) for both participant groups in all AOIs ... 186

Table 5.4 Summary of three-way ANOVA on entry time for all AOIs ... 187

Table 5.5 Summary of three-way ANOVA on dwell time for all AOIs ... 187

Table 5.6 Descriptive statistics of first fixation duration (ms) for both participant groups in all AOIs ... 188

Table 5.7 Descriptive statistics of number of revisits for both participant groups in all AOIs ... 189

Table 5.8 Summary of three-way ANOVA on first fixation duration for all AOIs ... 191

Table 5.9 Summary of three-way ANOVA on number of revisits for all AOIs ... 191

Table 5.10 Descriptive statistics of number of fixations for both participant groups in all AOIs 192

Table 5.11 Descriptive statistics of average fixation duration (ms) for both participant groups in all AOIs 194

Table 5.12 Summary of three-way ANOVA on number of fixations for all AOIs 195

Table 5.13 Summary of three-way ANOVA on average fixation duration for all AOIs 195

Table 5.14 Descriptive statistics of first mouse click time for both participant groups 196

Table 5.15 Summary of three-way ANOVA on first mouse click time 196

Table 5.16 Descriptive statistics of task completion time for both participant groups 197

Table 5.17 Summary of three-way ANOVA on task completion times 197

Table 5.18 Descriptive statistics of ratings on ease of use, ease of remembering and ease of learning for participants' assigned menus combination for both participant groups 198

Table 5.19 Summary of three-way ANOVAs on ratings of ease of use, ease of remembering and ease of learning for participants' assigned menu combination 200

Table 5.20 Descriptive statistics of ratings on ease of use, ease of remembering and ease of learning for all menu combinations for both participant groups 201

Table 5.21 Summary of three-way mixed ANOVAs on ratings of ease of use, ease of remembering and ease of learning for all menus combinations 203

Table 5.22 Descriptive statistics of ratings on ease of use, ease of remembering and ease of learning of menu combinations for both occasions 204

Table 6.1 Summary of methods for the three studies in this programme of research 217

Table 6.2 Summary of significant group level differences between non-dyslexic and dyslexic participants for eye gaze behaviour and performance 222

Table 6.3 Summary of significant differences due to text and menu presentation manipulations of non-dyslexic and dyslexic participants 223

Table 6.4 Summary of significant differences for preferences and opinions of non-dyslexic and dyslexic participants 224

Table 6.5 Recommendations made for people with dyslexia based on the text presentation and web navigation studies 226

ACKNOWLEDGEMENTS

I would like to extend my appreciation and gratitude to my supervisor, Professor Helen Petrie for her patience, motivation, support and knowledge. She has guided me during my years of study at the University of York and gave me countless of her valuable inputs.

Besides my supervisor, I would like to thank my thesis committee, Dr. Alistair Edwards and Dr. John Rack. Thank you for your insightful comments and encouragement, but also the hard questions during the viva which had helped me to widen my view in this research.

My family for supporting me spiritually throughout this journey and my life in general. Thank you for making me feel so close when in reality it is 6, 572 miles away. Thanks to the FaceTime!

I would also thank my colleagues and friends in the university for sharing their pearls of wisdom during this journey, for listening to all my rants and for giving impromptu inspirations to keep my enthusiasm in finishing what I have started.

Not to forget, participants who have participated in my studies. Without them, I would not be able to finish the studies. Thank you for the feedbacks that I received during the sessions.

Finally, my gratitude goes to Majlis Amanah Rakyat (MARA), a Malaysian government agency whom supported this research and my study at the university.

Above all, I owe it all to Almighty God for granting me the wisdom, health and strength to undertake this journey and enabling me to its completion.

DECLARATION

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

CHAPTER 1: INTRODUCTION TO RESEARCH AND DYSLEXIA

1.1 Introduction

According to the World Health Organization, WHO (2015), 15% people of the world's population have some sort of disability. The number of people with a disability is continuing to increase due to aging and chronic health conditions. As reported by International Dyslexia Association, IDA (n.d.), 15% - 20% of world's population have some symptoms of dyslexia. Dyslexia occurs in all parts of the world and is not related to intelligence.

Investigations of dyslexia started in 1887 with slow progress until the 1960s. In the beginning, research focused on finding causes, establishing theories, definitions, diagnosis and intervention or treatment plans for dyslexia (Washburn, Binks - Cantrell, & Joshi, 2014). Most of the research on the amelioration of difficulties associated with dyslexia focused on reading from the printed page, however more recently attention has been given to interaction with computers for people with dyslexia (McCarthy & Swierenga, 2010). According to de Santana et al. (2012), accessibility barriers exist for dyslexic people in using the web. However, research about dyslexia and web accessibility is sparse compared to that for other groups of disabled people such as blind people (Rello, Kanvinde, & Baeza-Yates, 2012a). Consequently, research for this dissertation will focus on how the digital adaptation of text presentation and web navigation can be used to help people with dyslexia.

Reading is one of the concerns in dyslexia, therefore this chapter is going to give a brief introduction on the word recognition process in reading in order to have better explanation on the problems had by people with dyslexia. The rest of this chapter will present an overview of dyslexia, its definition and related theories, and the characteristics of dyslexia. There is some controversy in the literature on exactly how dyslexia should be defined and whether people with dyslexia differ from other poor readers. This controversy will also be highlighted. This chapter ends with the structure of this dissertation.

1.2 Motivation for this Research

The Internet was intended to be used by everyone regardless of any disability (W3C WAI, 2018) and part of it is the World Wide Web (WWW), now simply known as the Web. Today, around 46% (around 3.5 billion) of the world population have access to the Internet (Internet Live Statistics, 2018). According to WebAIM (2016), creating web sites without web accessibility in mind may restrict the opportunity of some populations such as people with dyslexia to access the content. Therefore, web designers could increase accessibility through better designs that can be tailored to the needs of users with different requirements. With this in mind, the aim of the Web may not be achieved if documents and materials on the web are not well presented to people with dyslexia. In addition, the benefits of presenting documents in a manner that is optimised for people with dyslexia can also possibly benefit non-dyslexic readers as well (Zarach, 2002) since characteristics of dyslexia can also be found in non-dyslexic people in varying degrees (Rello, 2014a). However, Grigorovich-Barsky & Belson (2013) stated that documents designed for dyslexic people are not noticeably different from those designed for individual without reading disabilities. With the increment of Internet usage in current trend, not only the presentation of the material on the Web is crucial but how to navigate to the web resources is also important. According to Al-Wabil et al. (2007), highly textual content and poor navigation structure can burden people with dyslexia. Therefore, the main motivation for this research is to investigate how to improve the presentation of material on the Web and navigating inside Web to improve its access for people with dyslexia and if there is any difference between people with and without dyslexia.

1.3 Word Recognition Process in Reading

Reading is a very complex process of transforming print words to meaning and (if appropriate) speech. In order to understand dyslexia, it is important to understand how the brain works in processing print words. According to Coltheart (2005), most research on reading has agreed that there are two different approaches in transforming print words to speech and meaning. Figure 1.1 illustrates the dual-route approach which explains how printed words are processed during reading. The printed word is perceived by the eyes and the image is transmitted to the brain. Then, the word is visually analysed, to assess whether it is a familiar word or an unfamiliar word. If it is a familiar word, it will be processed using Route A, while unfamiliar words (new words or pseudo words) will be processed using Route B which involves decoding the orthography to phonology (Coltheart, 2005; Selikowitz, 1998; Siegel, 2006).

In Route A (known as the Lexical or Orthographic Route), the word is checked against the lexicon and meaning of the word can usually be rapidly retrieved from the semantic system. The meaning of the word and its pronunciation are stored in short-term memory while the speech generator activates appropriate parts of the brain to deliver the word for speaking, if needed. In Route B (known as the Non-lexical or Phonological Route), an unfamiliar word will be broken into segments such as graphemes, then are matched with sounds (phonemes) according to grapheme-phoneme correspondence rules and finally the sounds are blended together before the speech generator is initiated. Grapheme-phoneme correspondence rules are a set of relations between letter(s) which represent individual sounds (Kinsman, 1997). The overall process in this route is called phonological processing.

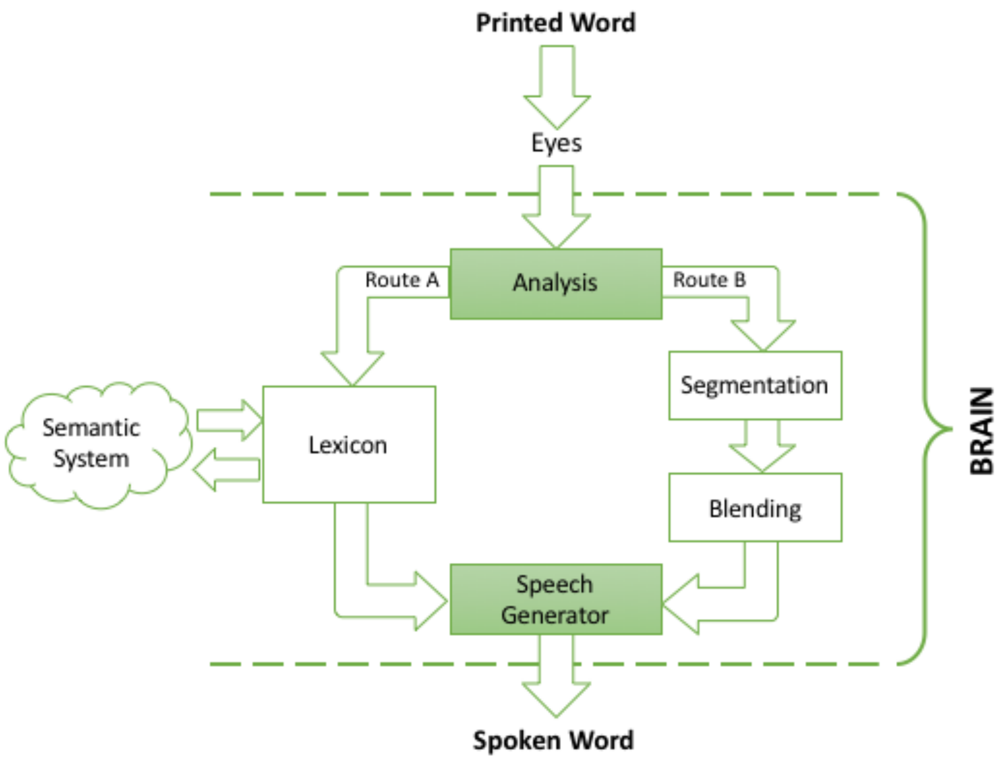


Figure 1.1 Dual-route model of reading (Source: Selikowitz, 1998)

According to Selikowitz (1998), children build up their lexicon as they learn to read. Two preparatory stages are involved before they can read automatically. A very young child starts with a *visual memory stage* where they will memorise the shape of the words. Eventually they need to override this stage as many words have similar shapes such as “Help!” and “Bang!”. Furthermore, the advancement of their spelling skills needs more than just visual memory. Later, the child will enter the *phonological stage* in which they will apply phonological processing during reading which helps in reading any word or non-word. In this stage, they learn to decode or break the words into segments and associate each sound with the segments. Eventually, they will not need phonological processing in reading and automatically access the lexicon whenever they encounter familiar words and read like an adult.

1.4 Dyslexia: Theories, Definition and Cause

In 1881, Oswald Berkhan first identified a person with dyslexia, a boy who had normal intelligence but who was unable to read (Snowling, 2000). Later in 1887, the ophthalmologist Rudolf Berlin introduced the term “dyslexia”. The word dyslexia comes from the Greek, meaning “difficulty with speech”[1]. Dyslexia is a disability that is characterised by difficulties of word recognition and decoding and poor spelling. It is considered to be a problem since it affects people’s ability in reading, writing and spelling (McLoughlin, Fitzgibbon, & Young, 1994).

Numerous theories and definitions of dyslexia have been suggested and discussed since the term was first discovered. According to Vellutino, Fletcher, Snowling, & Scanlon (2004), theories explaining the underlying causes of dyslexia can be divided into two broad categories of causes, the cognitive and the biological. Ramus et al. (2003) discussed five major theories explaining dyslexia:

(1) phonological theory – suggests deficits in language processing including grapheme-phoneme mental mapping, word recall, and/or retrieval of speech sounds
(2) rapid auditory processing theory – suggests difficulty in processing short and rapid acoustic stimuli
(3) visual theory – suggest impairment in processing information from letters and words from a written page
(4) cerebellar theory – claims deficit in the mild dysfunctional cerebellum part of the brain, which affects motor control, speech articulation and overlearned task automatization (such as driving and typing)

[1] http://www.oxforddictionaries.com/definition/english/dyslexia

(5) magnocellular theory – is a unifying theory that combined all the theories above. It suggests dysfunctional magnocellular part of the brain, which affects visual, auditory and tactile modalities

After decades of research, there are still lots of disagreement and debates on the evidence of the theories explaining the causes of dyslexia which can be seen from various definitions discussed in this section. Frith (1999) proposed a causal modelling framework where she proposed that the definition of dyslexia should involve three levels of description – behavioural, cognitive and biological. She emphasised the idea of defining dyslexia with only one of the levels of description will lead to paradoxes and incorrect diagnoses. Two extreme examples she illustrated were firstly that people who have poor reading skills (behaviour) without brain abnormality (biological) and no cognitive deficit (cognitive) are not dyslexic; secondly people with dyslexia who have successful reading remediation can have brain abnormality and cognitive deficit but show no problems in reading and writing. However, Frith (1999) did not propose any definition of dyslexia based on this framework.

Previous research also has suggested that it is important to consider the cultural context in constructing definition of dyslexia, as it can worsen the condition (Frith, 1999; Goulandris, 2003). For example, different writing systems can affect the diagnosis of dyslexia. This suggestion was supported by a case of a bilingual dyslexic boy who had reading and writing difficulties in English but not in Japanese (Wydell & Butterworth, 1999). In addition, research has shown that languages with different writing systems such as Chinese and English activate different parts of the brain (Tan, Laird, Li, & Fox, 2005).

Thomson & Watkins (1990) loosely defined dyslexia as:

> *"A severe difficulty with the written form of language independent of intellectual, cultural and emotional causation. It is characterised by the individuals' reading, writing and spelling attainments being well below the level expected, based on intelligence and chronological age."* (p. 3)

In addition to the characteristics defined by Thomson & Watkins (1990), Reid & Green (1996) include characteristics other than reading, writing and spelling in their definition as below:

"Dyslexia is a processing difference experienced by people of all ages, often characterised by difficulties in literacy, it can affect other cognitive areas such as memory, speed of processing, time management, co-ordination and directional aspects. There may be visual and phonological difficulties and there is usually some discrepancy in performances in different areas of learning. It is important that the individual differences and learning styles are acknowledged since these will affect outcomes of learning and assessment. It is also important to consider the learning and work context as the nature of the difficulties associated with dyslexia may well be more pronounced in some learning situations."

The British Dyslexia Association (BDA) has not established definition of their own but acknowledges various definitions of dyslexia and approved definition from Rose Report (British Dyslexia Association, 2018a). Sir Jim Rose and other experts in the field of psychology, reading and neuroscience produced a definition of dyslexia in the Rose Report (2009), a review of dyslexia commissioned by the United Kingdom government:

"Dyslexia is a learning difficulty that primarily affects the skills involved in accurate and fluent word reading and spelling. Characteristic features of dyslexia are difficulties in phonological awareness, verbal memory and verbal processing speed. Dyslexia occurs across the range of intellectual abilities. It is best thought of as a continuum, not a distinct category, and there are no clear cut-off points. Co-occurring difficulties may be seen in aspects of language, motor co-ordination, mental calculation, concentration and personal organisation, but these are not, by themselves, markers of dyslexia. A good indication of the severity and persistence of dyslexic difficulties can be gained by examining how the individual responds or has responded to well-founded intervention." (p. 30)

The range of definitions of dyslexia in the literature shows the lack of agreement in its characteristics and may have a number of causes. Because of the disagreement in the definition of dyslexia, the International Dyslexia Association (IDA) started a Definition Consensus Project in 1994 which developed a working definition. Later in 2002, the consensus group revised and expanded the 1994 working definition by adopting a definition of dyslexia from Lyon, Shaywitz, & Shaywitz (2003) :

"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."* (IDA, 2002)

The group then revisited the definition at a conference in 2016 and decided to remain with this definition by summing up *"no compelling reason to change the definition of dyslexia. The definition remains meaningful for research and practice."* (Dickman, 2017). From the definition, IDA (2002) acknowledges that phonological deficit might be the cause for difficulties having by dyslexia. In a systematic review of research on dyslexia, consensus on supporting the phonological theory was accentuated (Rack, Snowling, & Olson, 1992; Shaywitz & Shaywitz, 2005; Vellutino et al., 2004). Prior research has presented sufficient evidence that supports the phonological theory of dyslexia where dyslexics were significantly poorer than non-dyslexics in phonological tests (Ramus et al., 2003; Snowling, 1998, 2001). According to the phonological theory, people with dyslexia have a core cognitive deficit in the ability to break the words into segments, mapping the segments to the appropriate sound and retrieving the information to articulate that sound. This process requires awareness to connect the letters (orthography) to the sound (phonology) appropriately. The impairment of this whole process leads to poor foundation of reading alphabetic systems (Lyon et al., 2003; Michail, 2010; Ramus, 2001; Ramus et al., 2003; Shaywitz & Shaywitz, 2005; Siegel, 2006; Snowling, 2001). Referring to the word recognition process in reading described in previous the sub-section, children with dyslexia have difficulty to progress from the visual memory stage to the phonological stage due to this phonological deficit (Snowling, 2001). It is assumed that dysfunctions of certain brain areas are the cause for this deficit, a part of brain that is used to connect phonological and orthographic representations (Henry, Beeson, Stark, & Rapcsak, 2007; Ramus et al., 2003).

1.5 Characteristics of Dyslexia

According to Frith (1999), one of the challenges in describing dyslexia is the diversity of ways in which it is expressed in different people. In other words, dyslexia is a condition in which people show varying symptoms from individual to individual with different levels of severity. They show inconsistent results in different dyslexia assessments (Rello, Kanvinde, & Baeza-Yates, 2012b; Riddick, Farmer, & Sterling, 1997; Stanovich, 1996). It appears to be hereditary, has no relationship with level of intelligent and may co-morbid of other learning disabilities such as dyspraxia, dyscalculia, obsessive compulsive disorder (OCD), Tourette's syndrome, ADHD, specific reading disability (SRD), speech sound disorder (SSD) and specific language impairment (SLI) (Boada, Willcutt, & Pennington, 2012; Germanò, Gagliano, & Curatolo, 2010; Pauc, 2005; Siegel, 2006; Snowling, 2012; Washburn et al., 2014).

Dyslexia is a life-long condition, which causes difficulties in reading, writing and spelling. By adulthood, many people with dyslexia will have developed strategies in order to compensate for their difficulties to be able to undertake daily tasks in education or the workplace (Turner, 1997). The definition from IDA (2002) outlines key characteristics of dyslexia such as difficulties to recognise familiar word accurately, difficulties to read familiar word fluently, lack of ability to decode a word into its segmentation of sound thus blending the sounds together and also having poor spelling. Having some of the characteristics can lead secondary characteristics such as problems in reading comprehension and also reduced reading experience which hinder vocabulary and knowledge. The BDA (2018a) acknowledges characteristics of dyslexia as having difficulties in visual and auditory processing, slower processing speed of the brain, difficulties in recalling and naming objects, short term memory deficits, and a mismatch of automatic development with individual's other cognitive abilities (Malpas, 2012).

Apart from the key characteristics described in the definitions of dyslexia, people with dyslexia may have difficulty in writing with letter reversals ('b' for 'd'), letter omissions ('emty' for 'empty') and letter additions ('arround' for 'around') (Al-Wabil, Zaphiris, & Wilson, 2006; Gregor, Dickinson, Macaffer, & Andreasen, 2003; Ndombo, Ojo, Osunmakinde, & Phasha, 2013; Rello, Bayarri, Otal, & Pielot, 2014; Schattka, Radach, & Huber, 2010). They also may have difficulties with spelling such as making phonological errors ('f' and 'ph'), replacing letters with similar sound ('s' and 'z'), using wrong word endings ('ie' and 'y'), letters being written out of sequence ('does' and 'dose'), and confusion or omission of vowels and double consonants (Reid, 2011; Vidyasagar & Pammer, 2010). They also tend to write with inconsistent use of capital and lower case letters and tend to avoid writing lengthy texts (Reid, 2011).

In reading, some reported that people with dyslexia may have poor comprehension. People with severe dyslexia usually have difficulty this difficulty because they are struggling with decoding words rather than deriving their meaning (Snowling, 2000). In addition to poor comprehension, they may read at slower pace thus show reluctance to read for pleasure. They also tend to get lost in the line of text during reading and tend to have difficulty in the recognition of words, numbers, letters and punctuation. They may have difficulty in recognising and remembering the sound of words and tend to replace words with similar meanings when reading out loud. Moreover, people with dyslexia tend to confuse words that have similar outline shapes. For example, "either" and "enter" are both words that not only have the same first and last character but also have a similar overall shape (McLoughlin et al., 1994; Reid, 2011; Selikowitz, 1998; Snowling, 2000).

In speaking, some people with dyslexia may have difficulty pronouncing multi-syllabic words and difficulty in expressing their thoughts in words. They also tend to have poor awareness of rhymes (Davis, 1992; Reid, 2011).

1.6 The Prevalence of Dyslexia

Approximately 5% - 10% of people in the world have dyslexia (Vidyasagar & Pammer, 2010; Washburn et al., 2014). For the UK, the estimate is 10% of the population and 4% are severely dyslexic (British Dyslexia Association, 2018a; National Health Service, 2014). For English in the USA, the Interagency Committee on Learning Disabilities (1987) estimated that 10% to 17.5% of the population have dyslexia. The discrepancy of the prevalence rates for two countries using the same language is probably caused by cultural and social factors and the fact that figures are higher in urban populations (Snowling, 1998, 2013). Dyslexia is also found in other languages, but the prevalence estimates vary depending on orthographic depth of the languages (Brunswick, 2010; Vellutino et al., 2004). For other languages, estimates are 8.6% - 11% for Spanish (Rello & Barbosa, 2013), and 5% - 10% for French (APEDA-France, 2010). As discussed in previous section, there are a lot of disagreements on the definitions, characteristics, causes and theories related to dyslexia. It is not clear whether these estimates of the prevalence of dyslexia are based on which dyslexia manifestation. Furthermore, no information is provided in these references on the diagnostic tools used in the estimation, which is important as different tools will have different formulations. As prior research has suggested, the prevalence rates varied depending on which definition is used in identifying dyslexia, the criteria, cut-off point used in the diagnostic tools and whether the data are taken from clinical or large population samples (Katusic, Colligan, Barbaresi, Schaid, & Jacobsen, 2001; Rose, 2009; Snowling, 1998, 2013). Katusic et al. (2001) estimated the prevalence varied from 5.3% to 11.8% while Snowling (2013) estimated between 3% to 10%.

Spencer (2000) and Rello & Baeza-Yates (2000) noted that speakers of *transparent* languages have little difficulty in decoding written words. Transparent languages have a one-to-one relationship between graphemes (letters) and phonemes (sounds), and the spelling of words is very consistent in term of how they sound. On the other hand, English has a more complex orthography in which the same letter combinations produce different sounds and different letter combinations produce the same sound (Vellutino et al., 2004). This effect is illustrated in Figure 1.2, which shows same letter combinations coloured in blue and their corresponding pronunciation in the International Phonetic Alphabet. This characteristic may contribute to a higher proportion of the population having dyslexia among native speakers of English. For readers who have difficulties, the main problem in reading with transparent languages is fluency of reading, while in non-transparent languages such as English the main difficulty is reading accuracy (Elliott & Grigorenko, 2014).

cough /kɒf/	through /θruː/	rough /rʌf/
their /ðɛː/	there /ðɛː,ðə/	they're /ðɛː,ˈðeɪə/

Figure 1.2 Example of inconsistent relationship between letters and sounds in English

1.7 The Controversy and Misconception Surrounding Dyslexia

The range of theories, causes and characteristics to be included in the definition of dyslexia has led to controversy about how to differentiate people with dyslexia from people with other problems such as poor readers, people with dyspraxia, dyscalculia and other conditions. According to Dickman (2017), this controversy involved people who believed the definitions of dyslexia are too broad or too narrow and to what extent of range of characteristics should be included. By considering the controversy, it is unclear whether the percentages of the prevalence of dyslexia are either too low or too high. While this argument can be considered out of scope for this thesis, I considered it important to mention this controversy as an important issue that has received much publicity recently.

Elliott and Gibbs (2008) argued that including in the concept of dyslexia characteristics such as problems with memory and movement is too broad and oversimplified, thus failing to differentiate dyslexia from other conditions. Beside reading, writing and spelling problems, these characteristics are found in other disorders such as ADHD, dyscalculia and dyspraxia (Jeffries & Everatt, 2004; Washburn et al., 2014).

In these controversies, some of the characteristics that are disputed that part of dyslexia include they tend to be clumsy, uncoordinated and poor body balance because of dysfunctional in certain part of their brain (Michail, 2010). However, Elliott and Grigorenko (2014) argued that these difficulties might be caused by dyspraxia which can co-occur with dyslexia. Not only is there controversy about the inclusion of particular characteristics from other learning disabilities in the definition of dyslexia, pre-service teachers from the US and the UK often have misconceptions about dyslexia. Pre-service teachers are students who enrolled in a teacher preparation program and working toward teacher certificate. In a survey of pre-services teachers by Washburn et al. (2014), found that misconceptions about dyslexia included poor visual perception as a cause, gender imbalance in incidence and dyslexia being curable over time. An online survey by National Center for Learning Disabilities (NCLD) in the USA on public perceptions reported some misconceptions about learning disabilities (LD) including dyslexia (Cortiella & Horowitz, 2014). The survey which involved random samples of 1, 980 adults found that up to one third of the respondents had inaccurate perceptions on the causes of LD, 22% of them believed that LD is caused by watching too much television, 24% of them believed that LD is caused by childhood vaccinations and 31% of them believed that LD is caused by poor diet. In addition, more than one third of the respondents thought that lack of involvement from parents or teachers during childhood can cause LD.

Despite these disagreements in the definition of the condition, Elliott and Grigorenko (2014) stated it is widely agreed that difficulty in phonological coding is the main characteristic of dyslexia (Elliott & Grigorenko, 2014; Marinus et al., 2016; Snowling, 2000; Vellutino et al., 2004; Washburn et al., 2014; Wery & Diliberto, 2017). For the purposes of this programme of research, the definition of dyslexia from the IDA (2002) will be used in order to plan for user studies.

1.8 Aims and Scope of Research

The theories described in Section 1.4 attempt to explain dyslexia, its causes and effects. However, this programme of research will focus on practical implementations of how to facilitate access to information for people with dyslexia, particularly on how to best text to people with dyslexia. According to Schoonewelle (2013), there are two pillars in dealing with dyslexia, remediating and facilitating. In remediating, emphasis is given on training people with dyslexia in linguistic subskills. The focal point in facilitating is through methods to assist people with dyslexia. Facilitating accessibility for people with dyslexia will create equal opportunities for them in using web. Therefore, this programme of research will concentrate on the factors of how digital adaptation of text presentation and web navigation can be used to facilitate people with dyslexia. Text presentation and navigation are important, as five out of thirteen guidelines are associated with these aspects in the Web Content Accessibility Guidelines, WCAG 2.1 (W3C, 2019):

- Guideline 1.1 Text Alternatives: *Provide text alternatives for any non-text content so that it can be change into other forms people need, such as large print, braille, speech, symbols or simple language*.
- Guideline 1.4 Distinguishable: *Make it easier for users to see and hear content including separating information in the foreground from the background*.
- Guideline 2.4 Navigable: *Provide ways to help users navigate, find content and determine where they are*.
- Guideline 3.1 Readable: *Make text content readable and understandable*.
- Guideline 3.2 Predictable: *Make web pages appear and operate in predictable web*.

This programme of research on digital text presentation and navigation to support people with dyslexia aims to seek answers for the following research questions. The two main research questions are:

1. To what extent does the text presentation affect eye gaze behaviour, reading performance, preferences and opinions of adults with dyslexia compared to adults without dyslexia?
2. To what extent does the design of navigation menus affect eye gaze behaviour, navigation performance and opinions of adults with dyslexia compared to adults without dyslexia?

In understanding behaviour, this research is interested to obtain unbiased, objective and quantifiable data. Since eye movements are also influenced by visual presentation, eye tracking devices were used in this research to get insight of participants' behaviour and also performance. In addition, eye tracking devices also offer visual representation on how people have interacted with the presented stimuli in the research.

In term of text presentation, many aspects could be considered. However, to investigate all of the aspects simultaneously is complicated (Kuster, van Weerdenburg, Gompel, & Bosman, 2017). Therefore, particular aspects were chosen to investigate, these can be seen in the specific research questions. Therefore, the main research questions can be broken down into the following more specific questions:

1 (a) To what extent does the typeface used in the presentation of text on a computer screen affect eye gaze behaviour, reading performance, preferences and opinions of adults with dyslexia compared to adults without dyslexia?

1 (b) To what extent does the font size used in the presentation of text on a computer screen affect eye gaze behaviour, reading performance, preferences and opinions of adults with dyslexia compared to adults without dyslexia?

1 (c) To what extent does the line spacing used in the presentation of text on a computer screen affect eye gaze behaviour, reading performance and preferences of adults with dyslexia compared to adults without dyslexia?

1 (d) To what extent does the line length used in the presentation of text on a computer screen affect eye gaze behaviour, reading performance and preferences of adults with dyslexia compared to adults without dyslexia?

2 (a) To what extent does menu organisation in web navigation affect eye gaze behaviour, navigation performance and opinions of adults with dyslexia compared to adults to without dyslexia?

2 (b) To what extent does menu visibility in web navigation affect eye gaze behaviour, navigation performance and opinions of adults with dyslexia compared to adults to without dyslexia?

Another aspect of interest in this research are whether there are any differences between adults with different severity levels of dyslexia. Since people with dyslexia shows varying symptoms with different levels of severity, it is interesting to investigate whether adaptation of text presentation is due to different severity levels of dyslexia. This lead to the following specific questions:

3 (a) To what extent will adults with mild dyslexia have more efficient eye gaze behaviour, better reading performance and better ease of read on

dyslexia-optimised typefaces, in comparison to adults with moderate dyslexia

3 (b) To what extent will adults with mild dyslexia have more efficient eye gaze behaviour, better reading performance and better ease of read on larger text, in comparison to adults with moderate dyslexia

3 (c) To what extent will adults with mild dyslexia have more efficient eye gaze behaviour, better reading performance and better ease of read on wider line spacing, in comparison to adults with moderate dyslexia

3 (d) To what extent will adults with mild dyslexia have more efficient eye gaze behaviour, better reading performance and better ease of read on shorter line lengths, in comparison to adults with moderate dyslexia

The population of people to be used in this research is adults (aged 18 to 60) with and without dyslexia. Participants with dyslexia were categorised into different severity levels of dyslexia; mild and moderate dyslexia based on a diagnostic checklist developed by Snowling, Dawes, Nash, & Hulme (2012). All participants were native speakers of English and required to have normal or vision correctable with spectacles. The control groups of adults without dyslexia were matched with the dyslexic group for age, gender, computer experience, and educational level.

1.9 Structure of the Thesis

This thesis is structured as follows:

- Chapter 2 describes the literature review of relevant research. It covers text presentation, particularly on computer screens, and its impact on people with dyslexia, how they navigate in web sites, guidelines for text presentation for people with dyslexia, technologies available for people with dyslexia, and finally difficulties encountered by people with dyslexia when using the web.

- Chapter 3 presents Study 1 which investigated the effects of font size and typeface on detailed reading behaviour, comprehension and eye gaze behaviour of adults with dyslexia and adults without dyslexia.

- Chapter 4 presents Study 2 which investigated the effects of line spacing and line length on detailed reading, comprehension and eye gaze behaviour of adults with dyslexia and adults without dyslexia.

- Chapter 5 presents Study 3 which investigated the effects of menu organisation of a web site on navigation behaviour, navigation performance and information retrieval between adults with dyslexia and adults without dyslexia.

- Chapter 6 presents the overall discussion of the programme of research including the contributions made by the research and its limitations. At the end of the chapter, recommendations of future research are also made.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

The aim of this chapter is to give an overview of research on digital text presentation and navigation support to people with dyslexia. The chapter is arranged into five themes: text presentation, navigation, guidelines and recommendations, current assistive technologies for people with dyslexia, and the difficulties encountered by people with dyslexia in using web sites.

2.2 Text Presentation and Typefaces Designed for People with Dyslexia

In this section, to situate the work on text presentation for people with dyslexia, the characteristics of typefaces (also known as font families) are explained in order to provide an understanding of how these typefaces differ. This was followed by an introduction to dyslexia-optimised typefaces and their characteristics.

2.2.1 Characteristics of Typefaces

In typography, a typeface (also known as font family) is a set of one or more characters that share common design features. Typefaces therefore consist of a collection of letters, numerals or symbols with particular characteristics (Lawler, 2006). A font is a typeface which has specific style, size and weight. For example, Times New Roman is a typeface, however Times New Roman Bold 12-point is a font. In addition, Times New Roman Italic 12-point size is a different font from the previous example. However, it is a common misconception that most people are referring typeface as font and often those two words are used interchangeably (probably caused by menu options in word processors which ask users to choose a font, when they should ask people to choose a typeface). The basic characteristics of typefaces are spacing (whether the typeface is proportional or mono-spaced); style of the typeface (whether it is Roman, Oblique or Italic); and the category of typeface (whether it is Serif or Sans Serif typeface).

In proportional typefaces, different characters occupy different width spaces, for example "l" is narrower compared to "w". While in mono-spaced typefaces, every character occupies same width space even for letter "`l`" and "`w`" (Gelderman, 1999).

Figure 2.1 shows a sample of text with proportional typeface in the upper line and mono-spaced typeface in the lower line.

The quick brown fox jumps over the lazy dog (Helvetica)

`The quick brown fox jumps over the lazy dog` (Courier)

Figure 2.1 Example of text with Proportional (Helvetica, 14pt) and Mono-spaced (Courier, 14pt) typefaces

In terms of style, Roman type is the normal upright font presentation, except Oblique and Italic typefaces are slanted. The only difference between both Oblique and Italic is their glyphs. Italic has more 'curvy' glyphs while Oblique has same upright glyphs as Roman, except it is slanted to the right. Figure 2.2 shows differences of glyphs between Roman in the upper line; Italic on the middle line and Oblique in the lower line.

The five boxing wizards jump quickly.
The five boxing wizards jump quickly.
The five boxing wizards jump quickly.

Figure 2.2 Example of text Roman, Italic and Oblique shapes (all Garamond typeface)

In terms of typeface category, Serif is any typeface with small decorative lines appearing at the end of strokes in the characters, while Sans Serif is any typeface without such small decorative lines (Lawler, 2006). Figure 2.3 shows a Sans Serif typeface on the upper line and a Serif typeface in the middle line. The bottom line of the figure shows the extra small lines in the Serif typeface (marked in red). Serif typefaces were claimed to make letters more distinguishable because they emphasise the shape of words. Despite the modern look of Sans Serif typefaces, some of characters look quite similar and this can cause confusion on mirror letters especially for people with dyslexia (Gelderman, 1999). Examples of mirror letters are b and d; p and q.

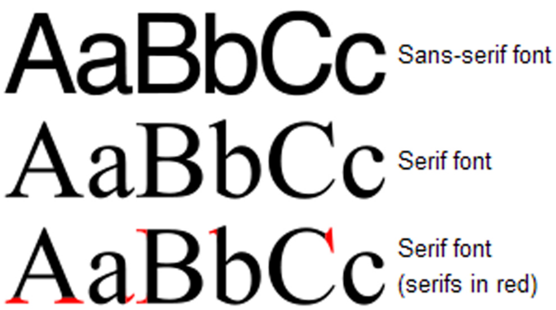


Figure 2.3 Example of Serif and Sans Serif categories of typefaces

2.2.2 Dyslexia-Optimised Typefaces

Designers have created special typefaces for people with dyslexia based on their difficulties during the reading process. Examples of dyslexia-optimised typefaces are Lexie Readable[2], Dyslexie[3] and OpenDyslexic[4]. Figure 2.4 illustrates sample of texts using the OpenDyslexic typeface on the middle line and Dyslexie typeface on the lower line, in comparison with the Arial typeface on the upper line.

The quick brown fox jumps over the lazy dog	(Arial)
The quick brown fox jumps over the lazy dog	(OpenDyslexic)
The quick brown fox jumps over the lazy dog	(Dyslexie)

Figure 2.4 Samples of text using Arial (14pt), OpenDyslexic (14pt) and Dyslexie (14pt) typefaces

Lexia Readable was designed by Keith Bates in 2004 (Bates, 2004) and later renamed to Lexie Readable. It has many similarities with Comic Sans MS without its "childish" appearance. According to K-Type (2008), Comic Sans MS has high legibility and helped lots of children with dyslexia to read, although they present no evidence to support this statement. Figure 2.5 illustrates sample of texts using Comic Sans MS typeface on the upper line and Lexie Readable typeface on the lower line.

The quick brown fox jumps over the lazy dog	(Comic Sans MS)
The quick brown fox jumps over the lazy dog	(Lexie Readable)

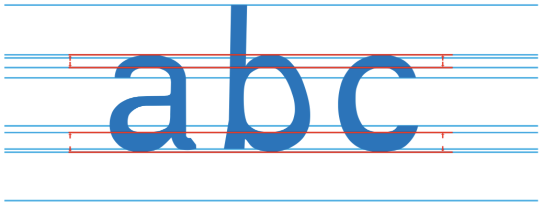
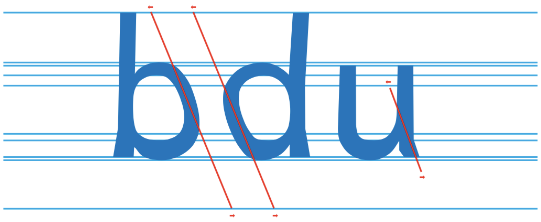
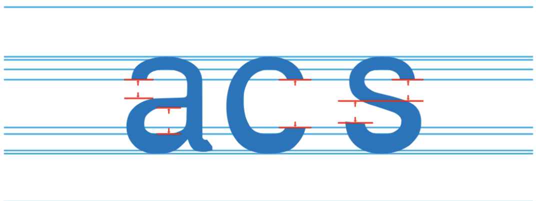
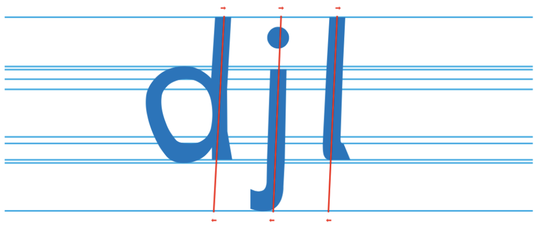
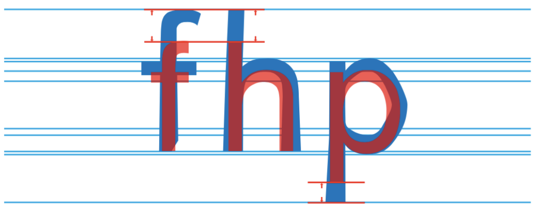
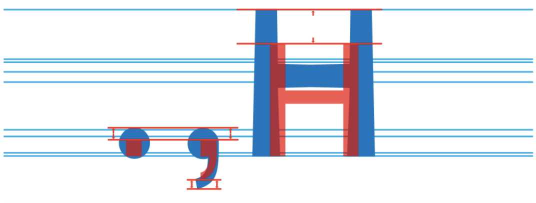
Figure 2.5 Sample of text in Comic Sans MS (14pt) and Lexie Readable typefaces (14pt)

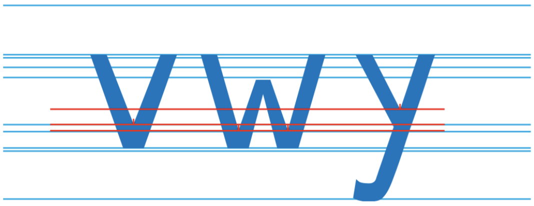
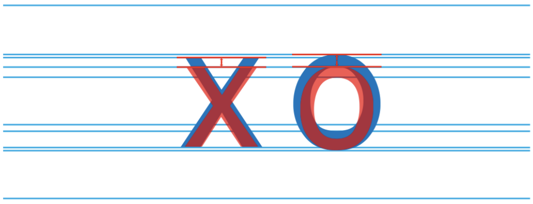
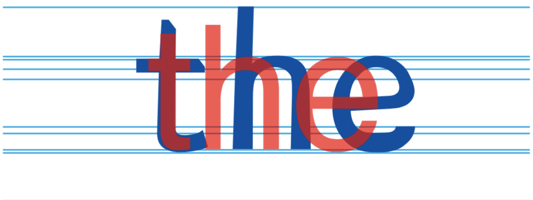
Dyslexie was designed by Christian Boer in 2008 as part of his PhD work (Boer, 2011). However, he did not conduct any empirical studies to measure the performance of dyslexics using Dyslexie as part of his thesis. The ideas underlying this typeface are that every character is unique and are intended to prevent readers from confusing them

[2] http://www.k-type.com/fonts/lexie-readable/
[3] http://www.dyslexiefont.com
[4] http://opendyslexic.org

(DyslexieFont.com, n.d.). This typeface comes with nine features as described in Table 2.1.

Table 2.1 Features of Dyslexie font (Source: DyslexieFont.com, n.d.)

Examples:	**Features Description**
	1. Heavy bottom: Lower character and heavier bottom act as "baseline" to prevent readers from seeing characters tipping upside down.
	2. Inclined letters: Changes in character's tail to avoid confusion with twin/mirrored characters such as 'b' and 'd'.
	3. Enlarged openings: Provide more space and makes characters more obvious.
	4. Slanted letters: Part of characters is semi-italic to prevent confusion with mirrored characters.
	5. Longer ascender and descender: It emphasised characters and helps to decrease switching letters.
	6. Bold capital letters and punctuations: Slightly bold capital letters and punctuation marks to help well in identifying start of sentences.

Examples:	Features Description
	7. Different heights: Helps similar looking characters resemble unique and different.
	8. Higher x-height: Spaces of characters between lines are increase while maintaining the width thus make it easier to recognise.
	9. Bigger spacing: With extra spacing between characters, words are standing out more clearly.

OpenDyslexic was designed by Abelardo Gonzalez as an open source typeface in 2011 (Gonzalez, 2015). Characteristics of this typeface include heavier weighted bottoms that indicate direction and reinforce the line of text, unique shapes of letters that are intended to help reduce confusion with letter reversals, and wide letter spacing that is intended to help readers from seeing the letters “r” and “n” when placed side by side as the letter “m” (OpenDyslexic, n.d.). Examples of these features in comparison with other typefaces are shown in Figure 2.6. The first column shows the name of the typeface, followed by the letters ‘r’, ‘n’ and ‘m’ in lower case; letters ‘M’ and ‘W’ in upper case; letters ‘d’, ‘b’, ‘q’ and ‘p’ in lower case; and letter ‘l’ (lower case L), the number ‘1’, ‘I’ (upper case i) , ‘i’, ‘j’ and ‘J’. These letters and number were chosen for the figure because they are commonly confused by people with dyslexia.

Figure 2.6 Unique letter shapes and heavier bottom of OpenDyslexic (Source: DyslexieFont.com, n.d.)

2.2.3 Criticism on Dyslexia-Optimised Typefaces

Dyslexia-optimised typefaces have received interest and support from academic institutions, international bodies, web site developers and book publishers . However, there have also been many criticisms of the dyslexia-optimised typefaces.

Kuster et al. (2017) questioned Boer's typographical design of the Dyslexie typeface. In typography, two measures are often used to define the legibility of typefaces: the ratio between lower case x-height to body height (the distance between the top of the tallest letter form to the bottom of the lowest one); and ratio between lower case x-height to capital height (the height of capital letters that are flat—such as H or I). However, the ratio of lower case x-height of Dyslexie is smaller than Arial and Times New Roman typefaces as illustrated in Table 2.2, which probably less legible than standard typefaces.

Table 2.2 Sizes of Dyslexie, Times New Roman and Arial in 12-point size (Source: Kuster et al., 2017)

	Dyslexie	**Times New Roman**	**Arial**
Body height (in mm)	5.6	4.2	4.2
Capital height (in mm)	4.0	2.8	3.0
Lower case height (in mm)	2.5	1.9	2.2
Ratio between lower case x-height to body height (in %)	44.5	44.8	52.0
Ratio between lower case x-height to capital height (in %)	63	67	73

Boer claimed that the characters in Dyslexie have unique and distinct designs to avoid confusion while reading (DyslexieFont.com, n.d.). However, Marinus et al. (2016) found that Dyslexie letters were visually less distinct than Arial. They measured this by computing the pixel overlap between lowercase and uppercase of each letter by aligning them to the central axis. In addition, Williams (as cited in Kuster et al., 2017) stated that distinct letters require more attention to be given to the visual shapes in comparison to the content of the text.

In order to find out whether it is the larger spacing or the Dyslexie typeface itself which improves reading performance of poor readers, Marinus et al. (2016) compared the default typographical design of Dyslexie with some adaptations of the Arial typeface. Table 2.3 shows the conditions used in this study, in which typefaces in all conditions had same x-height 0f 3.0 mm.

Table 2.3 Details of all conditions in a study with poor readers (Source: Marinus et al., 2016)

Condition	Typeface	Font size	Spacing	Example
1. Dyslexie	Dyslexie	14pt	Default Dyslexie spacing	The sky was dark.
2. Font size match	Arial	16pt	Default Arial spacing	The sky was dark.
3. Font size match and overall spacing match	Arial	16pt	Increased 1.5 points for between and within spacing	The sky was dark.
4. Font size match and balance spacing match	Arial	16pt	Increased 1.3 points for between words Increased 1.0 points for within words	The sky was dark.

Based on a study of 39 children with poor reading skills, Marinus et al. (2016) concluded that the Dyslexie typeface had no benefit in comparison to the Arial typeface with the same balance of larger spacing and height as the Dyslexie typeface. Indeed, spacing plays a bigger role on effective reading compared to the Dyslexie typeface itself. Reading with Dyslexie was 7% more efficient when the typeface had same x-height as

Arial. However, Dyslexie had no benefit when Arial had the same spacing with Dyslexie at the same x-height.

In other study on legibility and readability of dyslexia-optimised typefaces, Lexie Readable and OpenDyslexic, user performance and preferences were measured in the context of a voting application (Harley et al., 2013). User performance was measured based on event logs and eye tracking data, while subjective ratings on preferences, readability and legibility were rated in a post-study questionnaire. Seven participants with dyslexia (self-reported, aged 21.5 – 57.3) and five participants without dyslexia (aged 17 – 51.8) took part.

Using a voting application and an eye tracking device, participants were asked to read voting ballots, one in each of the dyslexia-optimised fonts and one in Helvetica, as a control condition. At the beginning, each participant was told which candidate to vote for on each ballot. After reading each ballot, participants rated it for legibility and readability of the typefaces and after reading all the ballots, they were asked to rank their preferences. There was no significant effect of typeface on the event logging and eye tracking measurements. In subjective measurements, Helvetica is preferred over both dyslexia-optimised typefaces (Lexie Readable and OpenDyslexic). This study concluded that OpenDyslexic did not lead to better reading performance and participants with dyslexia preferred the Helvetica typeface rather than dyslexia-optimised typefaces.

Based on these criticisms, there is a possibility that typeface designers have create dyslexia-optimised typefaces based on misconceptions they have about dyslexia (Washburn et al., 2014; Wery & Diliberto, 2017). The common misconceptions are that dyslexia is caused by poor visual perception which leads them to have problems with letter reversal, upside-down letters and so on. While some dyslexics have these problems, dyslexia is fundamentally a language processing problem and some dyslexics do not have those problems. There is further criticism about dyslexia-optimised typefaces which can be found in other studies which will be discussed in Section 2.3.

2.3 Impact of Text Presentation on Reading Behaviour of People with Dyslexia

To present text involves the selection of font types, font sizes, spacing for the characters, lines and paragraphs, line length and text and background colours. Each of these aspects is discussed below including related studies with participants with and without dyslexia, and its impact on their reading. Some common variables used in measuring reading performance are reading speed or rate, reading time, reading fluency, reading accuracy, and comprehension. In 1879, first era of eye movement research in reading have started and discovered basic measurements for eye movement (Rayner, 1998). Common variables in eye movement research in reading are number of fixations, fixation duration, number of saccades and number of regressions. When we read, continuous eye movements are called saccades, and when the eyes stop at certain points in between the saccades are called fixations. When the eyes move backward, it is called regression. On average, fixations last around 200 to 250ms, saccades last around 20-50ms with each saccade is between 7 to 9 letters when reading in English and 10% - 15% of saccades made are regression (Ellis, 2016; Rayner, 1998). Not only that, different eye movement were found between read out loud and silent. According to Rayner et al. (2012), eye movement are the ultimate tool to interpret the process of silent reading.

2.3.1 Research on the Effects of Typefaces

A number of studies have investigated which typefaces might benefit people with dyslexia. A well-considered typeface is important to maintain readability, organisation, easy navigation and consistency of text (Meindertsma, 2016). de Leeuw (2010) conducted an experiment on reading speed and accuracy while reading comparing the Dyslexie and Arial typefaces. Both typefaces applied same optical height. Two groups of participants, 21 with dyslexia (aged 19 – 25) and 22 without dyslexia (aged 19 – 28) took one and two minutes standardised Dutch reading tests. The first test (EMT) was used to measure word reading fluency while the second test (KLEPEL) used to measure non-word reading skills.

Each participant was assigned to one condition out of four, where each condition had two reading tasks with one auditory task between them. There were two versions of the EMT and KLEPEL test, used in both reading tasks. In each reading task, participants took both the EMT and KLEPEL tests, either printed in Arial or Dyslexie typefaces. All reading tasks used read out loud method. In the auditory task, participants need to classify some words under three levels of white noise. However, there was no explanation on the purpose of the auditory task and it was not used in further analysis. After completing all tasks, participants completed questionnaire about the Dyslexie typeface.

The results of this experiment showed people, with or without dyslexia, did not significantly increase their reading speed and reading accuracy with Dyslexie typeface. However, there was a trend for people with dyslexia to make fewer substitution errors (where a letter/consonant/vowel is substituted) with Dyslexie compared to Arial. In contrast, they made more read guessing errors (where reader guessed the word without reading it well). In the post-study questionnaire, people with dyslexia reported positive experiences with Dyslexie but reported that they would not use the typeface for the submission of assessments. This finding shows that participants with dyslexia are concerned about peers and instructors regarding their acceptance of a dyslexia-optimised typeface. It is important to note that this finding has no statistical support.

An extensive survey on Dyslexie typeface with adults, pupils, teachers, parents and therapist reported 76.8% of participants would recommend Dyslexie to others (DyslexieFont.com, 2012; van de Vrugt & Ossen, 2012). However, it is not clear whether these results are from dyslexic participants only or include responses from parents, teachers and others.

Pijpker (2013) reported similar results in an experiment on reading with adaptations of typefaces and background colours. Two groups of participants, 22 participants with dyslexia and 42 participants without dyslexia, were classified based on their reading level. For participants with dyslexia, 13 were classified as having a lower reading level (chronological aged 10 – 11) and 9 as having a higher reading level (chronological aged 11 – 12); while for participants without dyslexia, 12 were classified as having a lower reading level (chronological aged 8 – 10) and 30 with higher reading level (chronological aged 10 – 12).

In order to classify participants into dyslexic or not, each participant completed a test of two parts, 'Continuous Naming' and 'Word Reading'. In the first part, four naming tasks of colours, figures, images and letters were involved. While in the second part, two reading tasks within one minute were involved. At the end of this test, scores were used to indicate whether the participants had dyslexia, had a risk of dyslexia or no dyslexia symptoms. In addition to this, no explanation was given about how participants were differentiated into the different reading levels.

For the main study, there were two sets of four different texts. These sets had different reading difficulty levels, and all four texts within a same set had similar Flesch-Kincaid Reading Ease (FKRE) scores. According to their reading level, all participants read aloud four texts with combination of Dyslexie and Arial typefaces on White and Yellow coloured paper. Presentations of text were different for each reading level where participants with higher reading level given longer text and smaller font size.

Measurements of reading speed and reading accuracy were taken. Reading accuracy was measured through types of error and total of each error type. The results of this experiment are illustrated in Figures 2.7 and 2.8. Figure 2.7 shows interaction effect between typeface and background colour with reading speed. There was no significant effect on speed with participants with dyslexia when combinations of colour-typeface were used. Figure 2.8 shows interaction effect between typeface and background colour with reading accuracy. There was no significant effect on accuracy with participants with dyslexia when combinations of colour-typeface were used. However, participants with dyslexia at the lower reading level made less “word-deletion” errors (a word is deleted when reading sentences) when using Yellow-Dyslexie combination. Therefore, this might happen because of random occurrences.

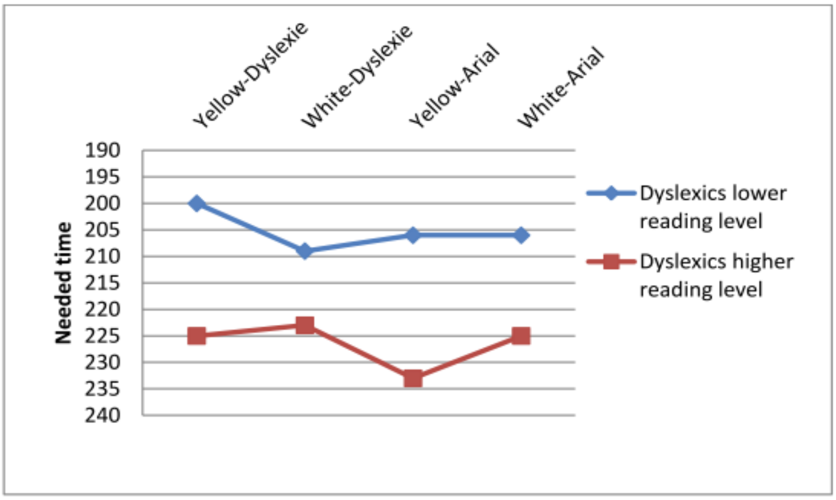


Figure 2.7 Effect of colour-typeface combinations on reading speed for participants with dyslexia (Source: Pijpker, 2013)

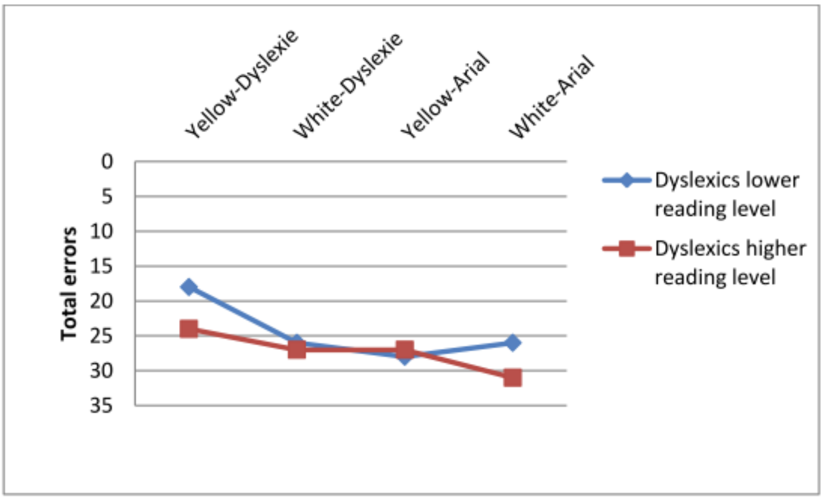


Figure 2.8 Effect of colour-typeface combinations on reading accuracy for participants with dyslexia (Source: Pijpker, 2013)

Another study involving Dyslexie and Arial typefaces is from Kuster et al. (2017). Two experiments were designed to answer questions whether Dyslexie lead to faster and accurate reading compared to Arial. In the first experiment, 170 children (100 males and 70 females) with dyslexia were included, aged 7.4 – 12.4 years. They were randomly assigned to two conditions, which had different sequence of typefaces. Participants for both conditions 'Arial-Dyslexie' and 'Dyslexie-Arial' had similar mean age and similar average Dutch reading skills.

Two version of cards were used where one card was printed in Arial (13-point size) and another card was printed with Dyslexie (12-point size, with some adjustment to line spacing to match with Arial). Both cards had the same sentences. The card contains short sentences with one to four syllables of words. Half of the participants were asked to read the card with the first sequence (either Arial or Dyslexie) and after second week, they were asked to read the card with another typeface in second sequence. Reading speed and accuracy were measured during both cards reading. Their preferences were asked in the second session.

For reading speed, the typeface read in the second time had significantly shorter time compared to the typeface read in the first time. The sequence of the typefaces did not have significant effect to reading speed and accuracy. Moreover, the number of errors recorded in second session did not decrease significantly from the typeface read in the first session. It was revealed that older participants read significantly faster and more accurate compared to younger participants. In further analysis, reading speed of Arial and Dyslexie did not differ significantly when read at different session. For preferences, regardless of the sequence of the typefaces, participants preferred Arial over Dyslexie.

With the same procedure, second experiment from the same study added Times New Roman typeface in the sequence of cards (Kuster et al., 2017). This experiment involved 102 children with dyslexia (Grades 2 to 6, aged 7.9 – 12.1 years) and 45 children without dyslexia (Grades 2 to 4, aged 7.6 – 11.9 years). Participants were randomly assigned in one of three conditions (Arial-Dyslexia-Times New Roman [ADT], Times New Roman-Arial-Dyslexie [TAD] and Dyslexie-Times New Roman-Arial [DTA]). Participants for each condition had similar age and Dutch reading skills.

Three cards with different difficulty were created. Card 1 contains words with one simple syllable, Card 2 contains words with one complex syllable which is hard to pronounce, and Card 3 contains multisyllabic words. In the other words, Card 1 is the easiest and Card 3 is the most difficult. Each card was printed into three versions of A, B and C where all versions had same words with different order and different typeface. Table 2.4 shows typefaces and font sizes used in the cards in all conditions. Both Card 1 and Card 2 had same x-height, however smaller x-height was used in Card 3. Some adjustments were made to interline spacing and body height of Times New Roman and Arial to match with Dyslexie which had bigger spacing and size.

Table 2.4 Typefaces and font sizes used in all three conditions (Source: Kuster et al., 2017)

Condition	Time	Typeface	Version	Font Size (in point)		
				Card 1	Card 2	Card 3
ADT	1	Arial	A	16	16	14
	2	Dyslexie	B	13	13	11
	3	Times New Roman	C	16	16	14
TAD	1	Times New Roman	A	16	16	14
	2	Arial	B	16	16	14
	3	Dyslexie	C	13	13	11
DTA	1	Dyslexie	A	13	13	11
	2	Times New Roman	B	16	16	14
	3	Arial	C	16	16	14

Note. 'ADT' indicates Arial-Dyslexie-Times New Roman, 'TAD' indicates Times New Roman-Arial-Dyslexie, 'DTA' indicates Dyslexie-Times New Roman-Arial.

Participants were asked to read all cards at three different session with one typeface at a time. They need to read out loud all the words in the card within one minute. The interval between each session was up to 2 weeks. After the third session, participants were asked to give their preferences on the fonts.

Participants without dyslexia read significantly faster than participants with dyslexia on all three cards. However, participants read Card 1 significantly more accurate than the other cards. There was no effect of typeface orders. For preferences, 33.3% of participants without dyslexia had no typeface preferences, 31.1% preferred Times New Roman, 22.2% preferred Arial and 13.3% preferred Dyslexie. For participants with dyslexia, 45.1% preferred Arial, 29.4% preferred Times New Roman, 13.7% had no typeface preferences and 10.8% preferred Dyslexie.

Both experiments shown that Dyslexie did not facilitate reading performance of dyslexic. However, it is important to note that there might have repetition effects or practice effects as participants read same text in one or two weeks intervening. Both experiments also showed no preferences over Dyslexie. The findings also in line with previous studies by de Leeuw (2010) and Pijpker (2013) where Dyslexie did not increase speed and accuracy. In addition, all studies were focusing on Dutch with participants with dyslexia.

Wery & Diliberto (2017) conducted a study that includes OpenDyslexic, Arial and Times New Roman. Twelve children with dyslexia (aged 9.0 – 12.8) participated. They were required to read out loud three reading tasks on letter naming, word reading and non-sense word reading in one minute for each task. Three sets of probes were printed, one set for each reading task. Probes for the letter naming task contained a list of randomly ordered uppercase and lowercase letter, probes for the word reading task contained a list of one to two syllables real words, and probes for non-sense word reading task contained a list of non-real words. List of words in the tasks were printed randomly for all participants.

Each set of probes was printed into three fonts which is OpenDyslexic 10pt, Arial 12pt and Times New Roman 12pt where all of the fonts had similar font heights. Words in all set of probes were arranged into 3 columns, double spacing row with black text on white background. There was no significant effect of typeface on reading accuracy or reading speed. In addition, OpenDyslexic was reported as having no positive effect on either measurement. However, this study involved a small number of participants.

In order to measure the impact of different fonts on reading performance and preferences for people with dyslexia, Rello & Baeza-Yates (2013, 2016) investigated 12 different typefaces (Arial, Arial Italic, Computer Modern Unicode, Courier, Garamond, Helvetica, Myriad, OpenDyslexic, OpenDyslexic Italic, Times, Times Italic and Verdana).

These 12 typefaces were analysed into five comparisons, where each typeface was categorised in one or more comparison groups:

1. Italic (Arial Italic, OpenDyslexic Italic, Times Italic) versus non-italic, roman (Arial, OpenDyslexic, Times)
2. Serif (Computer Modern Unicode, Garamond, Times) versus sans serif (Arial, Helvetica, Myriad, Verdana)
3. Mono space (Courier) versus proportional (Computer Modern Unicode, Garamond, Times)
4. Dyslexia-optimised (OpenDyslexic) versus non-dyslexia-optimised (Arial, Helvetica, Myriad, Verdana)
5. Italic dyslexia-optimised (OpenDyslexic Italic) versus italic non-dyslexia-optimised (Arial Italic)

Forty-eight participants with dyslexia (aged 11 – 50) and forty-nine participants without dyslexia (aged 11 – 54), took part in the study. In reading task, each participant read 12 texts of 60 words long. All 12 texts had similar genre; number of words; word length (4.92 – 5.58 characters) and contained no numerical expressions, acronyms or foreign words. Texts were presented with the same layout: left justified, 14pt font size, column width equal or less than 70 characters and displayed using black text on white background. Participants sat in a fixed chair and distance between participant and monitor screen was approximately 60cm. Dependent variables were number of fixations, fixation duration and reading speed which were recorded using an eye-tracking device. After reading task, comprehension was measured with multiple-choice questions and user liking was rated with 5-point Likert items.

For reading time and number of fixations for dyslexic group, there were no significant differences between italic, non-italic, serif, sans serif, proportional, monospaced, dyslexia-optimised, non-dyslexia optimised, italic dyslexia-optimised and italic non-dyslexia-optimised. Participants with dyslexia also had significant longer fixation duration with italic than roman; significant longer fixation duration with sans serif than serif; and significant longer fixation duration with proportional than monospaced.

Overall, participants with and without dyslexia had significantly slower reading times and longer fixation durations when reading with italic. It was observed that OpenDyslexic typeface did not lead to better reading performance. In term of participants' preferences, participants with dyslexia preferred Verdana and Helvetica, and disliked Garamond and OpenDyslexic. On top of it, participants without dyslexia preferred Verdana, Helvetica and Arial, and disliked OpenDyslexic in both roman and italic.

There are a number of methodological problems with this study. Participant age range too wide between children to adult and it were analysed together. In addition, authors did not specify value for line spacing in displaying text. As each typeface has different default value for line spacing, this might have an impact on participants' reading. Not only that, all the typefaces used had different x-height at the size of 14pt. These differences might have an effect on the results of eye behaviour and preferences. In addition, comparisons between typefaces are very confused with some typefaces appearing more than once, and it is unclear whether the authors are analysing the same data more than once.

Given above experiments, scarce research of typefaces was done on reading with adult dyslexic readers. It is noticed that typeface manipulations can help alleviate difficulties of people with dyslexia on reading ability. However, these experiments prove no significant effect on how specific typeface can help people with dyslexia and this include dyslexia-optimised typefaces as well. In addition to that, beneficial effects on certain typefaces had by adult participants may have partly occurred by typeface familiarity (Grigorovich-Barsky & Belson, 2013; Marinus et al., 2016).

2.3.2 Research on the Effects of Font Sizes

Not only the shape of typeface is important, font size also is crucial in determining legibility of texts. Instead on using typical font size, O'Brien et al. (2005) investigated the effect of Critical Print Size (CPS) on reading for children with dyslexia. The authors defined CPS as the minimum print size at which individual reads at their maximum speed. According to Legge & Bigelow (2011), reading with print size below CPS causes sharp declination in reading speed. Print size is measured as the height of a lower case of character 'x' (x-height). 34 children between 6.3 and 10.4 years were involved which had 22 children with dyslexia (mean age is 8.9 years old) and 12 children without dyslexia (mean age is 7.6 years old) took part. The children with dyslexia were on average younger than the children without dyslexia, as the two groups were matched for reading age. No participants had visual or hearing impairments.

Participants read sentences aloud from a flipbook as quickly as possible. The texts were in black text on white background in Times-Roman typeface. Each page in the flipbook had one sentence, and each sentence had 60 characters including spaces formatted into three lines of left-right justified text. There were two to three trials per print size. The sentences were presented at viewing distance of 40cm. For print size, 13 levels of Logarithm of the Minimum Angle of Resolution (logMAR) print size range from -0.2 to 1.0 (refer to Appendix A) were used. During reading, times per each page and reading errors were recorded. LogMAR value usually applied in logMAR chart, which comprises rows of letters and used by vision scientist, ophthalmologist or clinician to measure clearness of vision. Figure 2.9 illustrates LogMAR chart used to measure vision at 13 feet/4 meters.

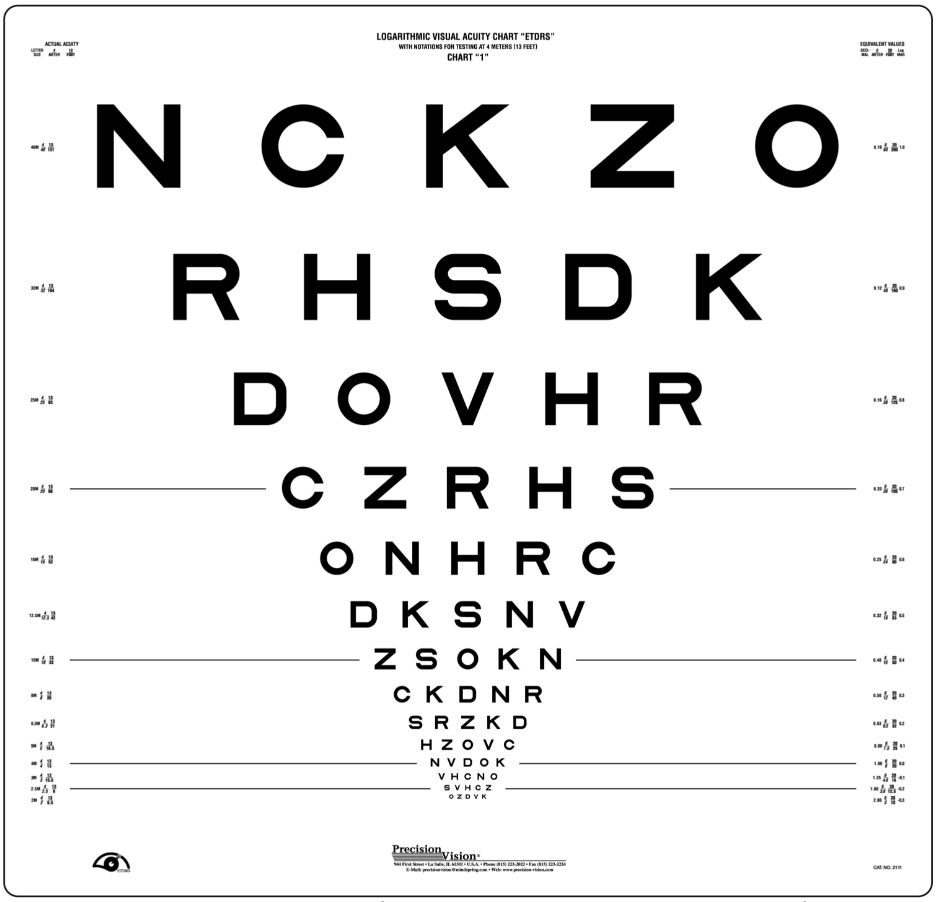


Figure 2.9 LogMAR chart (Source: Precision Vision, n.d.)

Result for reading speed by print size is illustrated in Figure 2.10. It shows example of individual reading speed for 2 non-dyslexic children and 2 dyslexic children. Both children with and without dyslexia were fit with two-limb function with maximum reading speed above critical print size (CPS) and their reading speed dropped below CPS.

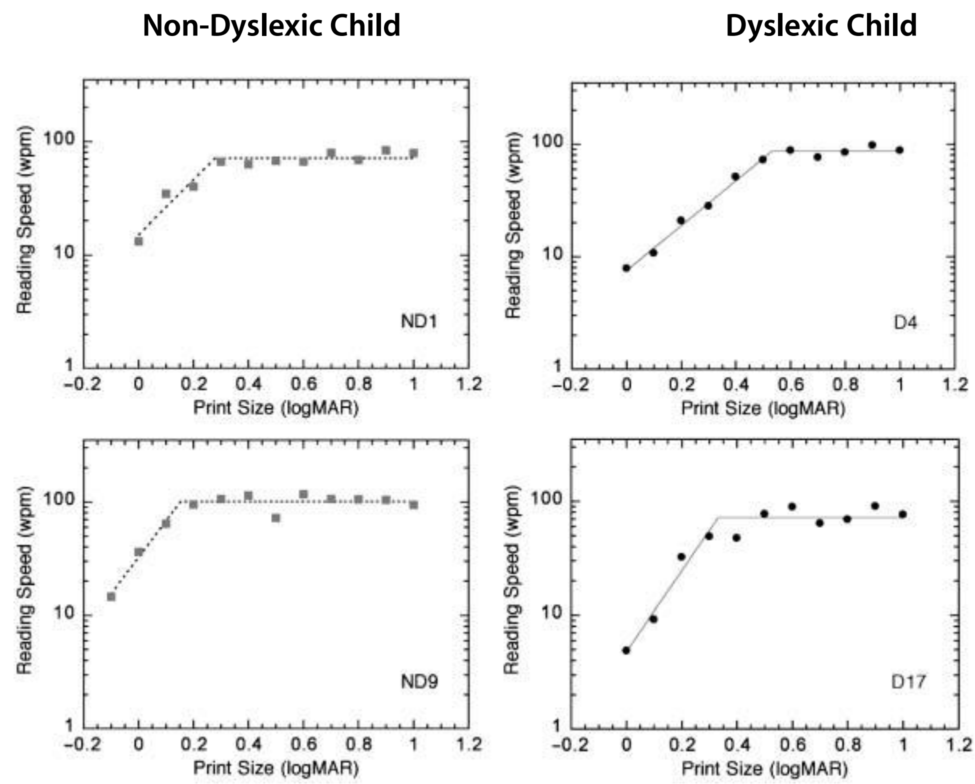


Figure 2.10 Example of reading speed for two children with and without dyslexia (Source: O'Brien et al., 2005)

In order to measure visual factors in reading, data for both groups were matched based on reading grade level and compared. Table 2.5 presented that children with dyslexia showed fastest reading speed with 32% larger print size compared to children without dyslexia that shared same chronological or reading grade level. It was pointed out that print size decreased with grade level where older children required smaller print size. However, no specific print size was recommended for people with dyslexia in this study.

Table 2.5 Means (Standard deviations) by group and reading grade equivalent for children with and without dyslexia (Source: O'Brien et al., 2005)

Reading Grade Equivalence	**Maximum Reading Speed (words per minute)**		**Critical Print Size (logMar value)**	
	Dyslexia M(SD)	**Non-Dyslexia M(SD)**	**Dyslexia M(SD)**	**Non-Dyslexia M(SD)**
1	23.34	44.88 (26.6)	0.452	0.295 (0.06)
2	81.46 (29.3)	102.56 (23.9)	0.265 (0.16)	0.177 (0.15)
3	118.16 (28.6)	112.43 (8.5)	0.168 (0.07)	0.098 (0.06)
4	117.25 (7.5)	163.31	0.100 (0.08)	0.028

A study by Rello, Pielot et al. (2013) focused on the readability and comprehension of texts of different font sizes and line spacing for people with dyslexia using eye tracking data. Twenty-eight participants with dyslexia (aged 14 – 38) involved, each read 24 Wikipedia articles. The 24 articles involved combinations of six levels of font sizes and four levels of line spacing. Font sizes chosen in this experiment were 10pt, 12pt, 14pt, 18pt, 22pt and 26pt; while line spacing were 0.8, 1.0, 1.4 and 1.8 lines. The authors had selected these values from previous recommendations and studies. The articles had similar characteristics, such as the same genre, topic, number of words (40 – 60 words per paragraph), paragraph structure, layout and few numerical expression, acronyms and foreign words.

Participants were asked to read only three paragraphs in the article. After each article, participants were asked to answer comprehension questions. Finally, they were asked to provide perception rating of each combination of font sizes and line spacing. The dependent variables were fixation duration, reading comprehension and participant perception (as rated on 5-point Likert items) of readability and comprehensibility. Distance between participants and screen presenting the text was fixed at 60cm.

Larger font sizes contribute to shorter fixation duration up to 18pt. Font sizes larger than that had no significant differences as illustrated in Figure 2.11. Figure 2.12 showed larger font sizes (14pt, 18pt, 22pt and 26pt) significantly produced higher comprehension scores than smaller font sizes (10pt and 12pt) for people with dyslexia. For participant perception, means of readability ratings significantly increased from font size 10pt to 18pt before it hit a plateau from 18pt to 26pt as illustrated in Figure 2.13. In addition, the means of comprehensibility ratings significantly increased from 10pt to 18pt and stabilised until 26pt as illustrated in Figure 2.14.

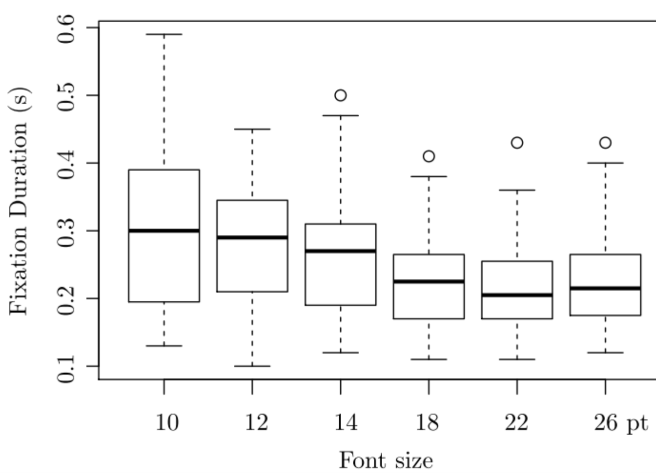


Figure 2.11 Fixation duration by font size (Source: Rello, Pielot, et al., 2013)

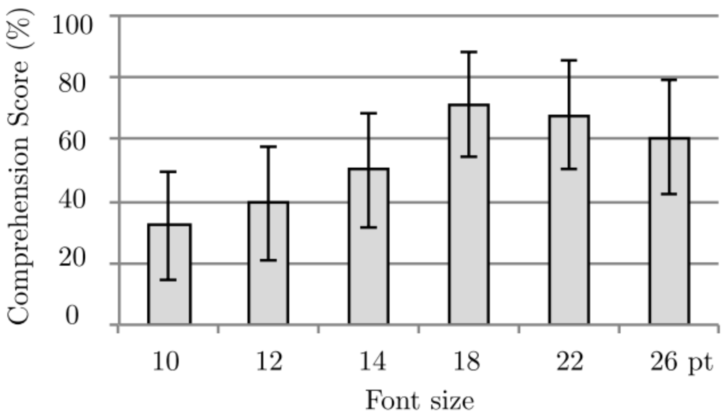


Figure 2.12 Comprehension score by font size (Source: Rello, Pielot, et al., 2013)

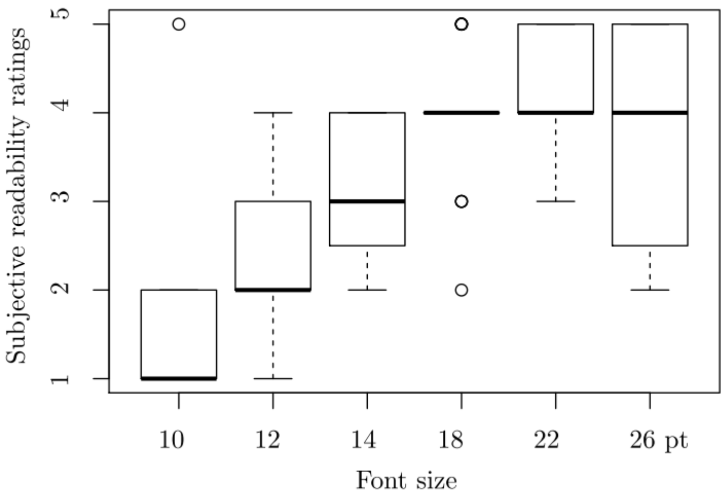


Figure 2.13 User rating for readability of font sizes (Source: Rello, Pielot, et al., 2013)

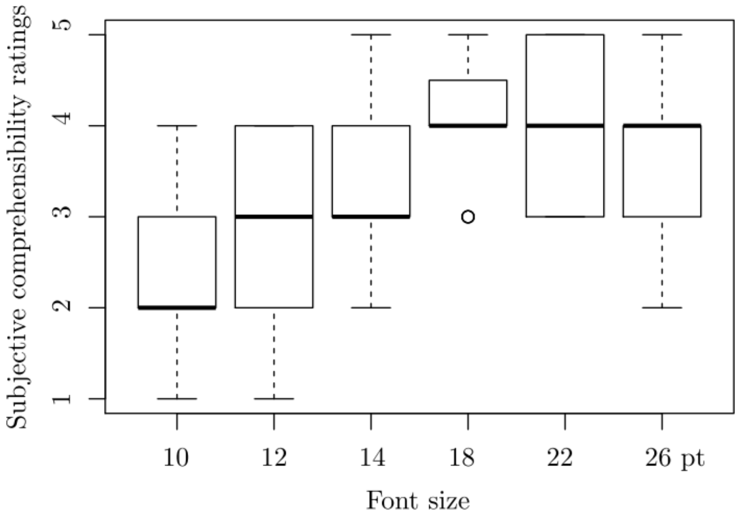


Figure 2.14 User rating for comprehensibility of font sizes. (Source: Rello, Pielot, et al., 2013)

On the basis of these results, Rello, Pielot et al. (2013) recommended the use of font size with 18pt for Web text and as they argued that this "strikes the balance between having the best readability, comprehension and subjection perception scores" (p. 6).

In spite of 18pt size, the recommended size was contradicted with 26pt in previous guidelines (Rello et al., 2012b). The guideline was based on a study conducted with 22 participants without dyslexia (aged 13 – 37) and 22 participants without dyslexia (aged 13 – 37). Later, the guideline was extended in another study by adding more participants and was slightly modified on the research design (Rello & Baeza-Yates, 2015).

The extended study of guidelines recruited 46 participants with dyslexia (aged 11 – 45) and 46 participants without dyslexia (aged 13 – 37). Participants were asked to silently read the text with an eye tracker and then were asked to rate their preferences. Figure 2.15 illustrates both groups of participants had shortest fixation duration and high preferences rating with 26pt. Overall, the study recommended size of 18pt – 26pt to be used in reading for both groups of participants. Both of this recommendation on optimal font size (Rello & Baeza-Yates, 2015; Rello et al., 2012b) for dyslexic and non-dyslexic are discussed in detail in Section 2.5.

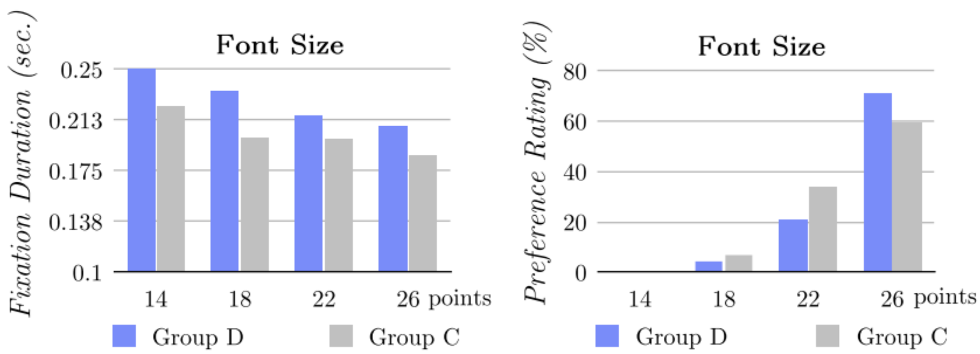


Figure 2.15 Reading performance and preferences of font size for participants with dyslexia (Group D) and without dyslexia (Group C) (Source: Rello & Baeza-Yates, 2015)

A study by Schoonewelle (2013) focused on comprehension and preferences of different font size, character spacing and line spacing for children with dyslexia. Thirty-nine participants with dyslexia (aged 12.4 – 15.6) involved, each reading eight texts in Dutch. The texts were taken from an establish digital test system which used to measure language, math proficiency and many other skills.

Two comparable groups were created, control and experiment groups. Both groups had comparable mean age and reading skills. Participants in control group read texts with default setting where the texts were displayed using font size 8.5pt (11px), character spacing of 0pt (0px) and line spacing of 1.0 (16px). Participants in the experiment group read texts with altered setting where font size 10pt (13px), character spacing of 0.5pt (1px) and line spacing of 1.15 (19px). Texts in both setting displayed in Verdana typeface, left-justified. Figure 2.16 and 2.17 illustrate both texts condition. All participants were given same amount of completion time to finish all the texts and its comprehension questions. They were asked to fill in a post-study questionnaire to indicate their preferences at the end of the study.

Beroemd worden

Het is een feit dat de jongeren van nu graag beroemd willen worden. Zij tonen overal op internet hun talent om maar ontdekt te worden. Volgens een onderzoek van www.mijnkindonline.nl en Mediawijzer wil 61% van de jongeren tussen de 11 en 17 jaar beroemd worden. Het is bijzonder dat ze in eerste instantie beroemd willen worden om andere mensen blij te maken. Pas ten tweede willen ze beroemd worden omdat ze heel rijk willen worden en een mooi huis willen hebben.

Om beroemd te worden, is het Internet natuurlijk een prima medium. Er zijn verschillende voorbeelden van jongens en meiden die het medium succesvol gebruikt hebben. Esmee Denters gebruikte YouTube, Lilly Allen gebruikte MySpace en Nicole Dee heeft Hyves gebruikt. Het succes van Esmee Denters heeft ervoor gezorgd dat er nu duizenden filmpjes op YouTube te vinden zijn van jongens en meiden die hopen ontdekt te worden. Maar wat is er nu allemaal voor nodig om echt beroemd te worden?

Figure 2.16 Text with default setting for control group (Source: Schoonewelle, 2013)

Beroemd worden

Het is een feit dat de jongeren van nu graag beroemd willen worden. Zij tonen overal op internet hun talent om maar ontdekt te worden. Volgens een onderzoek van www.mijnkindonline.nl en Mediawijzer wil 61% van de jongeren tussen de 11 en 17 jaar beroemd worden. Het is bijzonder dat ze in eerste instantie beroemd willen worden om andere mensen blij te maken. Pas ten tweede willen ze beroemd worden omdat ze heel rijk willen worden en een mooi huis willen hebben.

Om beroemd te worden, is het internet natuurlijk een prima medium. Er zijn verschillende voorbeelden van jongens en meiden die het medium succesvol gebruikt hebben. Esmee Denters gebruikte YouTube, Lilly Allen gebruikte MySpace en Nicole Dee heeft Hyves gebruikt. Het succes van Esmee Denters heeft ervoor gezorgd dat er nu duizenden filmpjes op YouTube te vinden zijn van jongens en meiden die hopen ontdekt te worden. Maar wat is er nu allemaal voor nodig om echt beroemd te worden?

Figure 2.17 Text with altered setting for experiment group (Source: Schoonewelle, 2013)

There was no significant difference on comprehension scores when texts were displayed using both default and altered setting. However, participants preferred altered setting which they found texts with bigger font size, larger character spacing and larger line spacing were easier to read.

It can be concluded there are benefits in manipulating font size to people with and without dyslexia. Based on experiments discussed above, O'Brien et al. (2005) did not provide any recommendation on ideal font size for people with dyslexia while 18pt to 26pt of font size were recommended (Rello & Baeza-Yates, 2015; Rello et al., 2012b). In spite of that, the size recommended were only tested with Arial typeface. Furthermore, different typeface might have different height for the same point size.

2.3.3 Research on the Effects of Line Lengths

Line length or column width is the distance between left and right edge of texts. According Dyson (2004, p. 379), *"line length can be measured by the physical length of the line"*. It is normally manipulated by visual angle, percentage of page width or characters per line (cpl). Schneps, Thomson, Sonnert et al. (2013) investigated the effects of device used in reading being held in the hand with different spacing. The texts displayed in the device had different line length, depending the size of the screen display. To investigate this, Apple iPad (PAD condition) and Apple iPod Touch (POD condition) were used either they were held in the hand while reading or not. In the POD condition, iPod has smaller screen display thus had shorter line length compared to PAD condition with Apple iPad. Texts displayed in both devices had conditions of normal spacing (NORMAL condition) or extra wide letter spacing (SPACED condition, increased by 29%). Each text had 208 words and were formatted using Georgia 32pt font, left-justified with 1.7 line spacing. In average, iPad displayed 11.6 words per line (67.2cpl) in one full page, while iPod displayed 2.19 words per line (12.7cpl) in twelve pages.

A group of participants with dyslexia, 27 high school students (aged 15 – 19) were recruited for this experiment. They were asked to complete a practice task, two reading tasks and recall session. In recall session, participants described the text they read, and their description were rated using 4-point scale (0 – unable to recall, 3 – recalled three or more details) for fidelity score. In the reading tasks, participants' eye gaze behaviour was recorded consisting of reading rate (RATE), number of fixations (FIX), inefficient saccades (TOT). Inefficient saccades contain backward, up and down eye movements.

The results of this experiment showed that by placing hand or not placing hand near the device had no significant effects on reading rate, number of fixations and inefficient saccades. Nonetheless of the combination of device, crowding and spacing, 90% of fidelity were scored at 3. There was also a significant effect on device where iPod had lower number of fixations and lower inefficient saccades.

Figure 2.18 shows participants with dyslexia read 27% faster, 11% less fixations and 40% less inefficient saccades when reading with iPod compared to iPad. In other words, participants were more efficient when reading text with shortened line length in iPod. Since both iPad and iPod had different screen size and displayed different line length in the experiment, it is unknown whether iPad with shortened lines will have similar positive effects on reading rate and fixations as in iPod.

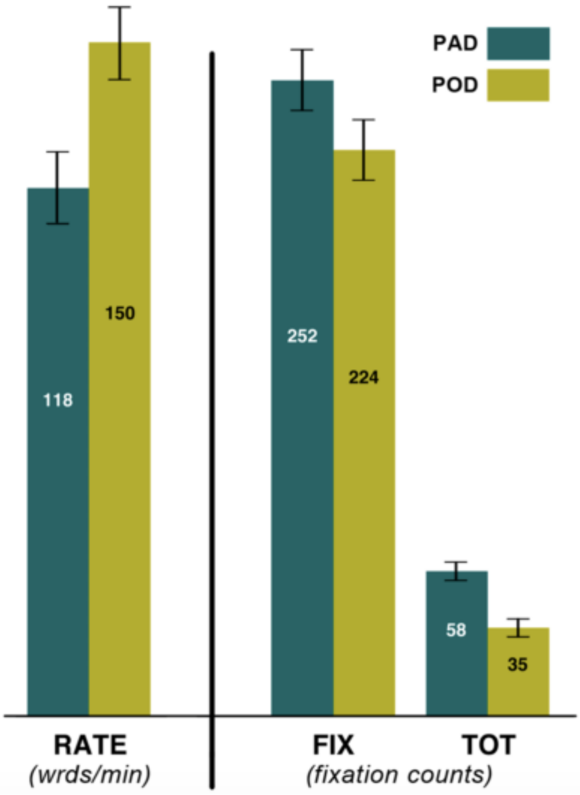


Figure 2.18 Reading with iPod was more efficient compared to iPad (Source: Schneps, Thomson, Sonnert, et al., 2013)

In an other study, line length consisting 77cpl was recommended in a layout guideline published for people with dyslexia (Rello et al., 2012b). Contradict with her recent guideline, no recommendation was made since no significant effect on fixation duration and preference rating were found for people with dyslexia (Rello & Baeza-Yates, 2015). For people without dyslexia, 44cpl were significantly preferred for reading as in Figure 2.19. The figure shows fixation duration and preference rating on column width (another term of line length) for both groups of participants. Both guidelines (Rello & Baeza-Yates, 2015; Rello et al., 2012b) were discussed in detail as in Section 2.5.

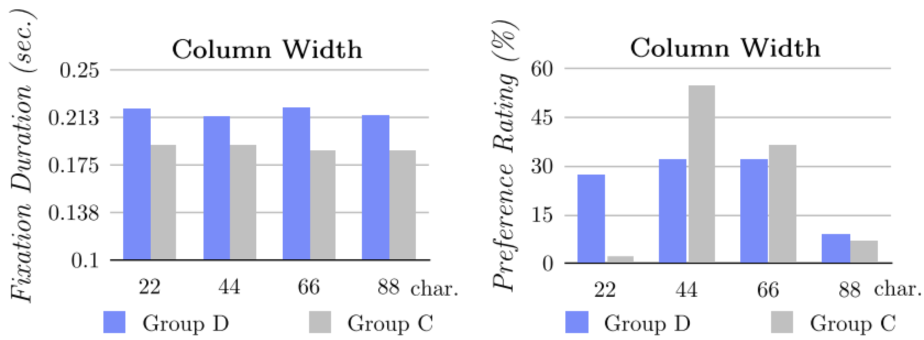


Figure 2.19 Reading performance and preferences of line length for participants with dyslexia (Group D) and without dyslexia (Group C) (Source: Rello & Baeza-Yates, 2015)

It can be seen that scarce research has been done on the effect of line length in reading. Not only that, there is no strong evidence on particular line length which can help people with dyslexia in reading.

2.3.4 Research on the Effects of Line Spacings

Another aspect of legibility in reading is spacing, where its effect on reading is being portrayed using term 'Crowding' (Montani, Facoetti, & Zorzi, 2014). According to Martelli et al. (2009, p. 2), crowding is an *"impaired recognition of a target due to the presence of neighbouring objects in the peripheral visual field"*. The term, crowding is referring to the inability to recognise objects in clutter therefore makes it difficult to identify objects or read (Levi, 2008; Whitney & Levi, 2011). Crowding takes place in the peripheral region which normal readers usually do not experience any difficulties. However for people with dyslexia, crowding not only affect their peripheral region, but also central vision (Zorzi et al., 2012).

As reading is influenced by crowding, it is important to avoid texts using space that is smaller than the critical spacing. Critical spacing is a distance where flanker letters (letters that distract reading target letter) starts to degrade reading performance. In addition, crowding worsens when multiple flankers are added and both target and flankers have similar contrast (for example, both target and flankers are black) (Levi, 2008; Spinelli, De Luca, Judica, & Zoccolotti, 2002). Crowding is said to influence the performance of reading in poor readers and people with dyslexia (Levi, 2008; Martelli et al., 2009). For an individual with dyslexia, crowding can have a negative impact on reading as they are more vulnerable during reading when letters are placed closely together (Marinus et al., 2016; Martelli et al., 2009).

Increments in spacing will lead to less crowding effect from neighbouring objects or letters therefore reading is fast (Levi, 2008). There are two different types of spacing in typography, horizontal spacing and vertical spacing. Letter spacing (or also known as intra-word spacing, character spacing, inter-letter spacing or tracking) and word spacing (or also known as inter-word spacing) are categorised in horizontal spacing. Line spacing (or also known as inter-line spacing, or leading) and paragraph spacing are categorised in vertical spacing. Several studies investigated how horizontal spacing can have an impact on reading with people with dyslexia (Martelli et al., 2009; Moll & Jones, 2013; Perea, Panadero, Moret-Tatay, & Gómez, 2012; Sjoblom, Eaton, & Stagg, 2016; Spinelli et al., 2002; Zorzi et al., 2012). Few studies were found for vertical spacing (Rello & Baeza-Yates, 2015; Rello et al., 2012b; Rello, Pielot, et al., 2013; Schoonewelle, 2013). Scarcity of research on vertical spacing with dyslexia shows that little attention has been given to this area (Bernard, Anne-Catherine, & Eric, 2007). This sub-section therefore will focus on research of line spacing with dyslexia. Research on other types of spacing are briefly discussed in this section and summarised in Table 2.15.

As discussed in section 2.3.2, Rello, Pielot et al. (2013) focused on the readability and comprehension of texts of different font sizes and line spacing for people with dyslexia using eye tracking data. The authors found no significant effect of line spacing with fixation duration as illustrated in Figure 2.20. For comprehension score, line spacing 0.8 was significantly higher as shown in Figure 2.21. There was no significant effect of line spacing on user perception on each readability and comprehensibility. Figure 2.22 shows fixation duration for bigger font size with all line spacing were decreased significantly. Lower fixation duration means better readability. Combination between 26pt font size and 1.4 line spacing had significant increment on fixation durations, indicates less readability.

Authors recommended default line spacing 1.0 because it did not have significant effect on reading experience of participants with dyslexia and most of web sites use this value in their HTML. In addition to this, 18pt size was recommended for people with dyslexia in this experiment as discussed in sub-section 2.3.2.

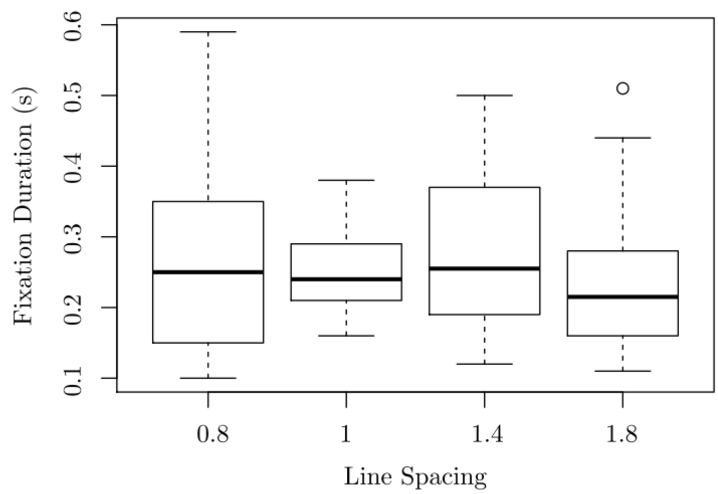


Figure 2.20 Fixation duration by line spacing (Source: Rello, Pielot, et al., 2013)

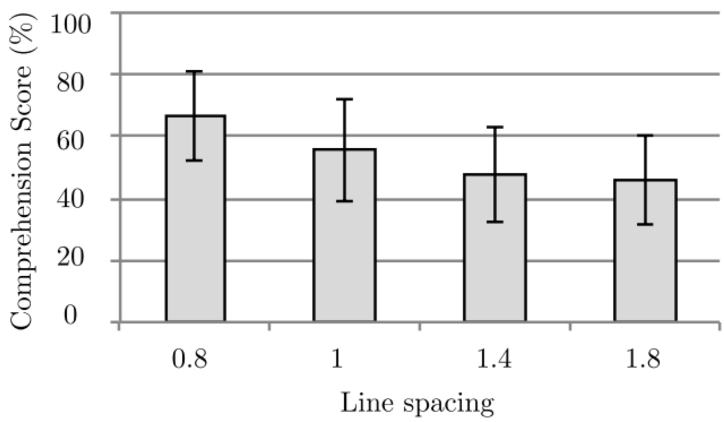


Figure 2.21 Comprehension score by line spacing (Source: Rello, Pielot, et al., 2013)

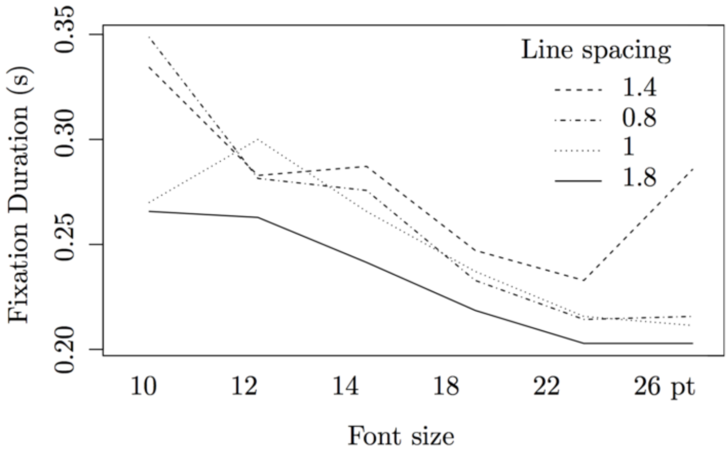


Figure 2.22 Interaction between font size and line spacing (Source: Rello, Pielot, et al., 2013)

However the recommendation of 1.0 line spacing contradicted her previous study and her recent guidelines (Rello & Baeza-Yates, 2015; Rello et al., 2012b). Line spacing of 1.4 was recommended in her previous study (Rello et al., 2012b), while no recommendation was made in her recent guidelines (Rello & Baeza-Yates, 2015). In the guidelines, line spacing had no significant main effect and no significant interaction between line spacing and group on both fixation duration and preference rating (Rello & Baeza-Yates, 2015). Figure 2.23 illustrates both groups of participants had similar fixation duration for all line spacing and high preferences rating between 1.0 and 1.2 line spacing.

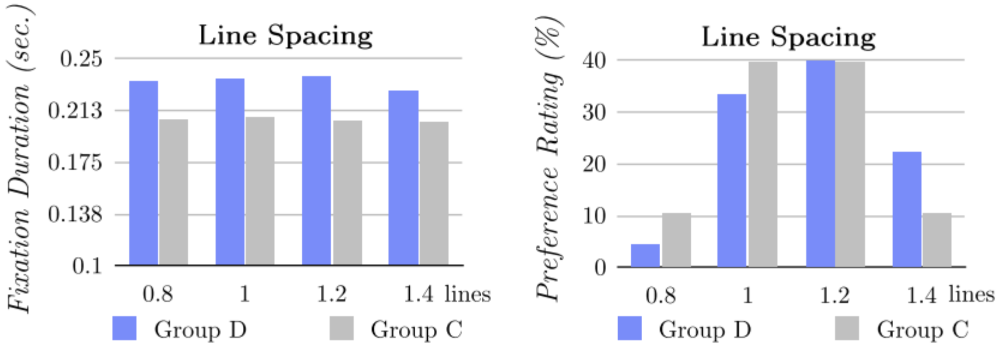


Figure 2.23 Reading performance and preferences of line spacing for participants with dyslexia (Group D) and without dyslexia (Group C) (Source: Rello & Baeza-Yates, 2015)

As discussed in section 2.3.2, Schoonewelle (2013) found that there was no significant difference on comprehension scores when texts were displayed using both default and altered setting. However, participants preferred altered setting which they found texts with bigger font size, larger character spacing and larger line spacing were easier to read.

For other types of spacing, study by Zorzi et al. (2012) measured crowding effects of dyslexia when reading using transparent and opaque languages. This study involved 34 Italian dyslexic children (to read in Italian, a transparent language), 40 French dyslexic children (to read in French, an opaque language) and 30 Italian non-dyslexic children. All of the participants were aged between 8 – 14 years old and had same reading level and IQ. They read 24 texts using Times 14pt on white A4 paper. All texts were left-justified.

Two conditions of character spacing were tested. Standard condition has normal character spacing (2.7pt) with default word spacing while spaced condition has bigger character spacing (5.2pt) with bigger word spacing (3 spaces). Different experiments were conducted in this study where line spacing were manipulated. In the first experiment, standard condition had default line spacing; while spaced condition had line spacing of 2.0. Figure 2.24 illustrates both standard and spaced condition in the first experiment. For second experiment, both standard and spaced condition had line spacing of 2.0. Second experiment was conducted to differentiate the effects of bigger character spacing from the line spacing. The dependent variables all experiments were reading speed and reading accuracy.

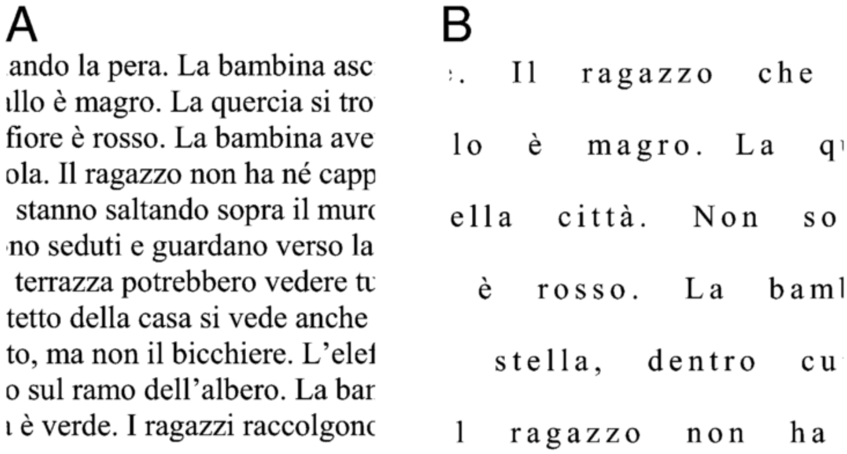


Figure 2.24 Sample of text read by participants. (A) Normal condition, (B) Spaced condition (Source: Zorzi et al., 2012)

This study found that extra-wide character spacing ameliorate reading ability (in terms of reading speed and accuracy) of participants with dyslexia more than wide line spacing in both transparent and opaque languages. This conclusion was partially supported (in terms of reading speed) by a study from Perea & Gomez (2012) where crowding on character spacing is found to influence reading speed of participants with and without dyslexia. In addition, reading speed is continued to decrease when character spacing exceeds some critical point (Perea & Gomez, 2012). However, Skottun & Skoyles (2012) disagree with Zorzi et al.'s (2012) finding by argued the significance testing used and a possibility of floor effect in reading accuracy for control group.

Schneps, Thomson, Sonnert et al., (2013) investigated the effects of character spacing by comparing iPad and iPod being held while reading. In both devices, texts were displayed with normal spacing or extra wide character spacing with an increment of 29%. 27 participants with dyslexia (aged 15 – 19) were recruited in this experiment. This experiment has been discussed in previous Section 2.3.3. For crowding, normal spacing had positive significant effects on reading rate and number of fixations. Contradict with Zorzi et al. (2012), where the participants with dyslexia showed significant increment in reading rate when reading with spaced condition in the first experiment.

Overall, studies with line spacing have shown that participants with dyslexia had no significant effect on fixation duration (Rello & Baeza-Yates, 2015; Rello, Pielot, et al., 2013) and comprehension (Schoonewelle, 2013). However Rello, Pielot, et al. (2013) reported that participants with dyslexia had higher scores with line spacing 0.8. In addition, studies found that bigger line spacing are easier to read (Schoonewelle, 2013) and line spacing 1.4 was preferred (Rello et al., 2012b). It can be seen that scarce research of line spacing has been done and confusing recommendations has been made. Furthermore, these studies only measured one variable for eye gaze behaviour, which is fixation duration (Rello & Baeza-Yates, 2015; Rello et al., 2012b; Rello, Pielot, et al., 2013).

2.3.5 Research on the Effects of Colours

According to McCarthy & Swierenga (2010), one of key problems encountered by people with dyslexia during web surfing is poor colour selection. There are little studies that focused solely on people with dyslexia with colour manipulation and it could be beneficial for reading using computer screen.

Rello & Baeza-Yates (2012) investigated which optimal colours can improve readability for people with dyslexia using eye-tracking data and user preferences. In this experiment, 23 participants with dyslexia (aged 13 – 37) and 92 participants without dyslexia (aged 13 – 43) were involved.

Eight colour combinations taken from previous literatures, recommendations, usage frequencies in Web and good ratio of luminosity contrast. Figure 2.25 shows details of colour combinations. Each colour pairs were applied to random comparable text with same word length (22 syllables), genre, rhythm, spacing (in terms of line, word and character) and layout (Arial 20pt font, left justified text). However, this experiment provided no information on spacing. Each participant read a set of text using eye-tracking device and answered a questionnaire on their preferences.

Text and Background Colour Pair (HEX code)	Colour Difference	Brightness Difference
Black (000000) & White (FFFFFF)	765	255
Black (000000) & Yellow (FFFF00)	510	226
Black (000000) & Cream (FAFAC8)	700	244
Off-Black (0A0A0A) & Off-White (FFFFE5)	735	245
Blue (00007D) & White (FFFFFF)	640	241
Dark Brown (1E1E00) & Light Green (B9B900)	310	137
Brown (282800) & Dark Green (A0A000)	240	107
Blue (00007D) &Yellow (FFFF00)	635	212

Figure 2.25 Combinations of colour pairs with colour and brightness differences (Source: Rello & Baeza-Yates, 2012)

Based on this experiment, 32.67% participants without dyslexia and 13.64% participants with dyslexia preferred Black & White colour pair. For colour pair Black & Yellow, participants with dyslexia showed largest fixation duration indicating longer time needed. However, majority of dyslexia participants preferred this combination. Black & Cream combination had the fastest reading speed when read by participants with dyslexia. None of dyslexic participants preferred Off-Black & Off-White. Dark Brown & Light Green and Brown & Dark Green showed high fixation duration mean and least preferred by participants with dyslexia as markedly by Checkpoint 2.2 in W3C Working Draft suggestion where brightness differences less than 125 and colour differences less than 500 are not recommended (W3C, 2000).

Figure 2.26 shows detailed result of this experiment. This experiment also pointed out no correlation between user performance and user preferences because word length of 20 syllables is too short to draw strong conclusion in this experiment. Furthermore, there was no statistical data presented in this experiment.

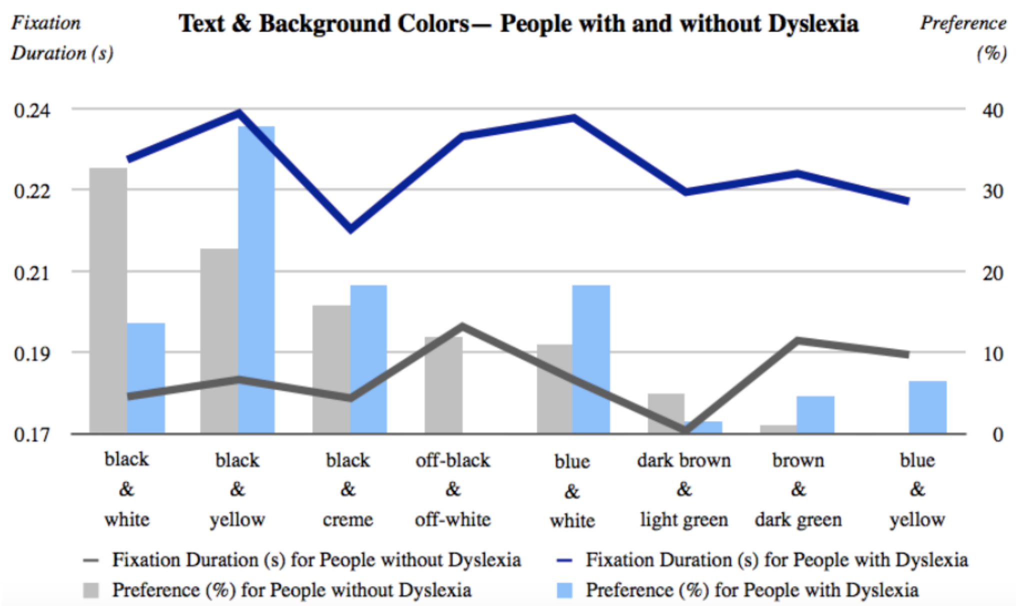


Figure 2.26 User performance and user preferences of colour pairs (Source: Rello & Baeza-Yates, 2012)

Rello and Bigham (2017) carried out an experiment to study the effect of background colours on screen readability. A total of 252 participants without dyslexia (male = 57 and female = 195) and 89 participants with dyslexia (male = 20 and female 69) involved in the experiment. Their ages ranged from 18 to 60.

Ten texts were created which had same genre, style and absence of numerical expressions, absence of acronyms and same number of words (55 words). All texts had different background colours, black Arial text at 18pt size and were left-justified. Figure 2.27 shows all ten colours with its HEX code and luminosity contrast ratio between background and text colour.

Blue:	#96ADFC	9.68:1	Purple:	#B987DC	7.56:1
Blue Grey:	#DBE1F1	16.05:1	Red:	#E0A6AA	10.20:1
Grey:	#D8D3D6	14.21:1	Peach:	#EDD1B0	14.35:1
Green:	#A8F29A	15.83:1	Orange:	#EDDD6E	15.17:1
Turquoise:	#A5F7E1	16.99:1	Yellow:	#F8FD89	19.40:1

Figure 2.27 HEX code and luminosity contrast ratio for all colour pairs (Source: Rello & Bigham, 2017)

This experiment was completed online where participants need to have laptop or desktop with mouse and Chrome browser installed. They were asked to complete a consent form and demographic questionnaire before reading instruction. They need to read ten texts in silence and complete a comprehension question for each text. While reading, reading time and mouse distance were measured. In order to track mouse movement, an open source software was used to log mouse movements at fixed-time intervals. Mouse distance represents the total number of pixels that the mouse travelled over the text. In addition, comprehension score was measured as control variable.

The difference in reading time for both groups was statistically significant where Peach background had the shortest reading time. Additionally, Orange and Yellow background colours had the second and third shortest reading time. However, Blue Grey background had the longest reading time. For mouse distance, participants with dyslexia had significantly longer mouse distance compared to participants without dyslexia. Long mouse distances could be related to difficulty in text readability. Not only that, the text with Blue Grey background also had the longest mean for mouse distance in both groups of participants. Overall, black text with warm background colours (such as Peach, Orange and Yellow) were recommended instead of black text with cool background colours (such as Blue Grey, Blue and Green).

Another experiment as previously discussed in Section 2.3.1, Pijpker (2013) examined if reading performance of children with and without dyslexia are positively affected by font type and colour combinations. Two coloured A4 papers (80 grams) to represent background colours, Yellow (CYMK (-0.01, 0.02, 0.31, 0.00)) and White; and also two typefaces, Arial and Dyslexie were used in this experiment. Both typefaces were printed using black colour on the coloured papers. Measurements for this experiment were reading speed and reading accuracy. As illustrated in Figure 2.7 and Figure 2.8, typeface and colour combination had no significant effect on reading speed and combination of Yellow and Dyslexie typeface produce higher accuracy for participants with dyslexia with lower reading level, while participants without dyslexia with higher reading level significantly made significant less "addition of letters" error when using Yellow coloured paper. However, there was no significant effect on accuracy and speed when combination of typeface and colour were used for participants with dyslexia.

In addition to the studies (Rello & Baeza-Yates, 2012; Rello & Bigham, 2017), black text with cream background colour was recommended in a layout guideline produced by Rello, Kanvinde, et al. (2012b). However in recent guideline (Rello & Baeza-Yates, 2015), the experiment on text/background with eight pairs of colours showed no significant effect on fixation duration and preference rating. Therefore, no recommendation on particular text and background colour were made and the selection of colours should be left to user. Figure 2.28 shows reading performance and preferences of both participants with and without dyslexia on colour combinations. Both guidelines were discussed in detail as in Section 2.5.

From the studies discussed above, it can be seen that previous researchers defined the name of same colour with different HEX code. For example, Yellow colour was represented with #FFFF00 (Rello & Baeza-Yates, 2012) and #F8FD89 (Rello & Bigham, 2017). However, W3schools.com stated when using Yellow as colour name in supported modern browsers, the colour will be displayed using #FFFF00 HEX code (W3schools.com, n.d.). In addition, there is no strong evidence on how specific combination of background and font colours can help people with dyslexia. Not only that, recommendation using colour names is confusing.

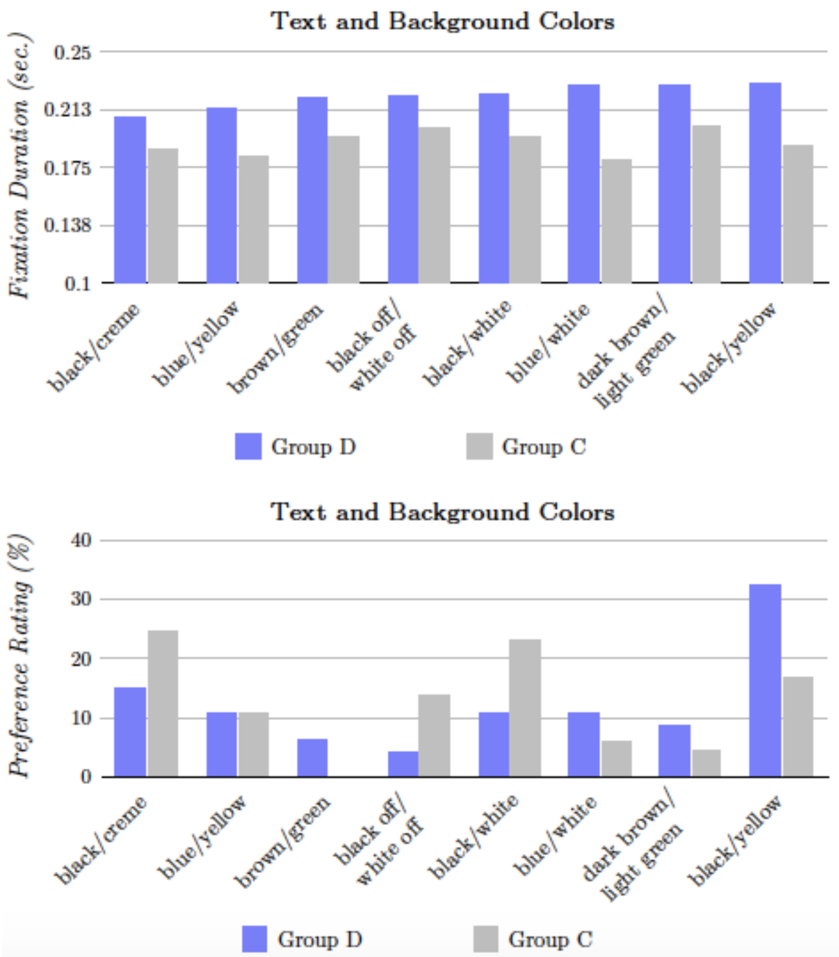


Figure 2.28 Reading performance and preferences of text/background colours for participants with dyslexia (Group D) and without dyslexia (Group C) (Source: Rello & Baeza-Yates, 2015)

2.4 The Way of People with Dyslexia Navigating Web Sites

Navigation is an important element in a website. According to Kalbach (2007), the term *web navigation* has different meanings depending on the contexts: (1) theory and practice of going from one to another, (2) process of goal-directed seeking and locating hyperlinked information, and (3) all user interface elements that provide access to pages. In general, navigation consist of user experiences and process of retrieving and locating information in a web. Highly textual content in the web and poor structure of navigation are burden to people with dyslexia (Al-Wabil et al., 2007). Therefore, it is essential to properly design navigation menu in order to reduce cognitive processing for people with dyslexia.

An exploratory study by Al-Wabil et al. (2007) focused on issues how navigation can reduce cognitive workload on working memory, sequencing and information organisation. All the three issues are part of problems had by people with dyslexia. In the study, ten participants with dyslexia (aged 18 – 49) were involved and some of them reported having other disabilities such as dyspraxia and sensory defensiveness. Sensory defensiveness is a condition of having negative reaction to sensations such as touch, movement, sound, texture or smell.

Semi-structured interview was conducted and participants' thoughts were gathered on four different topics consists of browser features, layout, site structure and navigation. Each topic was discussed using 15 slides and snapshots of websites with different examples were projected on whiteboard to each participant. For each slide, participants were asked questions and were encouraged to share their opinions and experiences.

This study identified four recurring behaviours in navigation which are (1) navigating within a web site and finding content on a web site, (2) keeping track of web sites and returning to previously visited web pages, (3) navigating within a web page and finding specific content on a web page, and (4) navigating the web and finding information across web sites. However, Al-Wabil et al. (2007) only discussed the navigation aids that emerged from first behaviour in detailed. Navigation aids discussed were sitemap, site index, search box, browser's back/forward button, navigation trails and menus.

Seven of participants did not use site maps because they were unaware of its existence or found it difficult to use. Heavy textual content and unsuitable structure were the reasons of their reluctance and they favoured site maps with hierarchical tree structures. Conversely, site index with alphabetically listing which have similar as with site map were used by nine of participants in finding information. However, their experience with site index was affected by their alphabet sequencing difficulties.

Out of ten participants, eight of them utilised the web site's search box as their last resort in finding information. This is because the quality of the search result was incomparable to Google Search and most local search function could not tolerate with their spelling mistakes. With a similar proportion, eight participants preferred to use browser's back/forward button rather than web site's back/forward button. Even though the website's back/forward button provides same functionality as browser's back/forward button, but from their perspectives it was different things and they were avoiding understanding the web site's back/forward button.

For navigation trails, six participants were familiar with the breadcrumb but only four of them found it useful. Four participants did not even notice the crumb because of the size used to present the trails was small. All participants agreed navigation menu should be consistent. However, some of the participants had bad experience with dynamic menus and preferred to have whole navigation bars visible at all times. For people with dyslexia, they found it difficult to control navigation which has sensitivity to cursor movements; hidden text appeared behind dynamic menus when hovering it and readability of text on semi-transparent menus. In brief, authors concluded despite the existence of many web sites conforming to accessibility standards, participants with dyslexia still have difficulties in navigation.

Despite of insights discovered by this study, there was a small number of participants were involved and findings of this study are based on reported experiences rather than direct observations on real web sites. Furthermore, there is no indication whether the participants' experiences are based on WCAG conformance or non-conformant web sites.

In order to gain better understanding on viewing behaviour using eye tracking data during web navigation, Al-Wabil et al. (2008) did another experiment. In this experiment, two participants with dyslexia and five participants without dyslexia were involved with nine tasks consists of six navigational tasks and three informational tasks within three conditions. Each participant looked for specific pages in navigational task and specific information in informational task. There were three WCAG compliant (EdenSkills, BBC, Tesco) and three non-compliant (eBay, Telegraph, Yahoo) web sites involved in both tasks. Measurements were number of fixations and fixation duration.

Findings on this experiment are shown in Table 2.6 and Table 2.7. Table 2.6 shows participants with dyslexia need longer time to complete all tasks compared to participants without dyslexia. Table 2.7 shows that participants with dyslexia exhibit more number of fixations and gaze time in contrast to non-dyslexic participants.

Table 2.6 Task completion time on three conditions for dyslexia and control group (Source: Al-Wabil et al., 2008)

	Experiment Conditions			
	1 – Navigational (WCAG conformance web sites)	**2 – Navigational (WCAG non-conformance web sites)**	**3 – Informational (WCAG conformance web sites)**	**All Tasks**
Dyslexia M(SD) (sec)	70.6 (46.5)	34.9 (4.5)	38.5 (7.8)	144.0
Non-dyslexia M(SD) (sec)	26.4 (4.4)	30.7 (3.4)	60.1 (11)	117.2

Table 2.7 Results on eye movement metrics between dyslexia and control group (Source: Al-Wabil et al., 2008)

Eye movement metric	Dyslexia	Control
Number of fixations (count)	332	304
Total fixation duration (sec)	134.7	115.6
Mean fixation duration (sec)	0.433	0.409

One of the tasks required participants to locate a target link in a web page. Figure 2.29 shows viewing patterns of participants with dyslexia appeared less strategic by showing more scattered region of interest. In contrast to participants without dyslexia, their eye gaze focused more on left side of navigational area. Figure 2.30 shows participants with dyslexia also are less confident in clicking target link because they need to revisit triple times to target area and double confirmation of the link before actually clicked on it.

Figure 2.29 Heat map showing region of interest for all participants (Source: Al-Wabil et al., 2008)

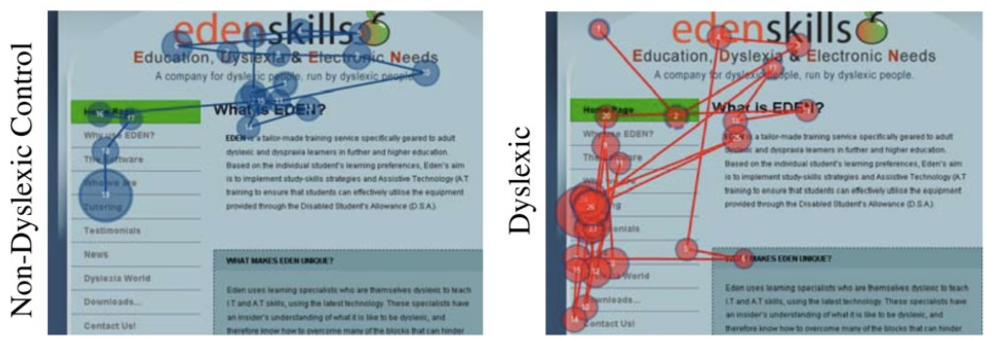


Figure 2.30 Gaze plot of fixation before selecting target link for one participant (Source: Al-Wabil et al., 2008)

Area of Interest (AOI) were examined and comparison between scan paths was calculated using Levenshtein distance. Levenshtein distance is an algorithm used in comparing similarity between two sequences. The greater the value of Levenshtein distance, the more different the two sequences are. In other words, one of the sequences require high number of changes to transform the sequence into the other. This algorithm has widely used in spell checking and DNA analysis. In this study, the authors defined scan path as visual attention formed by sequences of fixations and saccades. Then, Levenshtein distance was used to provide measure of similarity between these two scan paths. Scan paths taken from two dyslexia and two non-dyslexia participants. Table 2.8 shows distance between scan paths of two dyslexia and two non-dyslexia participants on three WCAG compliance web sites. Participants with dyslexia had higher mean of distance indicated they have more variability in scan paths.

Table 2.8 Results of Levenshtein distance for scan paths between participants with and without dyslexia (Source: Al-Wabil et al., 2008)

Levenshtein Distance		**Participants with Dyslexia**		**Participants without Dyslexia**	
		D1	**D2**	**ND1**	**ND2**
Stimuli	EdenSkills	7	6	6	16
	BBC	22	20	19	22
	Tesco	24	13	14	12
Mean (SD)		15.33 (3.3)		14.83 (2.6)	

This experiment concluded that participants with dyslexia required more cognitive processing in navigating a web site compared to participants without dyslexia. In addition, participants with dyslexia appeared to be less strategic during navigation and searching information. At the same time, limited numbers of participants with dyslexia is a significant weakness of this experiment since eye-tracking experiment requires some acceptable amount of participants in order to generate a reliable heat map or gaze plot to represent population of dyslexia (Pernice & Nielsen, 2009). Even though this experiment considered both WCAG compliance and non-compliance web sites, it is unknown if the navigation menu used in the web sites played any roles in the navigation strategy during navigation and searching information. This experiment also did not present any statistical analysis thus decrease its conciseness and clarity of interpretations.

Even though this study had considered both WCAG-compliance and non-compliance web sites, there is a possibility that organisation of menus influenced participants' eye gaze behaviour and performance.

2.5 Guidelines and Recommendations Published for People with Dyslexia

Web guidelines and recommendations explain how to make web content accessible to people with disabilities. Guidelines by Zarach (2002) stated 10 recommendations to improve web accessibility for people with dyslexia and also benefit people without dyslexia. The guidelines include customisable functionality for colours and fonts, recommendations to ensure web site easy to print out and read by screen readers, usage of images to replace long words, usage of numbered lists for point form information, short and simple text, consistent design throughout all web pages, usage of site map and straightforward navigation menu and the use of Sans Serif fonts with 12pt size. Author recommended using Verdana or Arial font types. However, author provides no scientific evidence on these guidelines.

Clear Text for All is a guidelines by Evett and Brown (2005) that combined Dyslexia Style Guide (DSG) from British Dyslexia Association (BDA) and Clear Print Booklet (CPb) from The Royal National Institute of the Blind (RNIB). Aim for these guidelines is to aid people with dyslexia and visually impaired readers to read clearly. Both guidelines of DSG and CPb were compared and unified into new guidelines. DSG were used as the main guidelines since recommendations from DSG covers wider aspects than CPb. In certain recommendations that partially exist in one of the guidelines, authors used their own judgment whether it is compatible for both target groups. Clear Text for All covers recommendations for text properties, colours, graphics, organisation and comprehension, words usage in content, and also screen readers. These guidelines provide no scientific evidence and no expert validation on their recommendations.

Guidelines by Friedman & Bryen (2007) focused for people with cognitive disabilities including dyslexia. Twenty web accessibility guidelines were chosen and compiled in a simplified list. In total, 187 separate recommendations were generated. Recommendations that overlapped and duplicated were removed from the list. After that, recommendations were ranked based on their percentage of how many times each recommendation was mentioned in the 20 guidelines. For recommendations that have been cited more than 15% were considered significant, therefore it was included in the new guidelines. This new guidelines consists of 22 recommendations that covered text properties, navigation, page design and layout, use of icons and pictures, writing style, margins, hyperlinks, and line spacing. Even though this set of guidelines was generated from 20 existing guidelines, there is a possibility of omission of relevant guidelines in literature review and authors did not explained why 15% citation were considered significant in their method. In addition to this, guidelines chosen were mixed up between cognitive disabilities, cognitive impairments, learning disabilities, dyslexia, aphasia, mental retardation and intellectual disabilities. Authors provide no scientific evaluation of these guidelines.

Dyslexia Style Guide (DSG) is a guidelines by British Dyslexia Association (2018b) and consists of five parts: font, heading and structure, colour, layout, and writing style. In the font part, it focused on text presentation including typeface, font size, character spacing, word spacing, line spacing and font style. In the heading and structure part, it focused on having consistent style for heading through MS Word ribbon and presentation of hyperlink. In the colour part, it focused on text and background colours, contrast levels and print material. In the layout part, it focused on text justification, line length and white spaces. In the writing style part, this guidelines focused on variety suggestion on writing and visuals.

Guidelines by de Santana et al. (2012) consists of recommendations that are linked to responsibilities of web site stakeholders such as the developer, the designer and the content producer. In these guidelines, recommendations were taken from existing guidelines and relocated into few groups. Table 2.9 shows a mapping between recommendations and level of involvement with each stakeholder's roles. These mappings served as dependency of guidelines on developer, designer and content producer. It was set to 'High' if the guidelines group depends solely on responsibilities of certain roles, 'Medium' if the guidelines group depends highly on certain roles and the others, and 'Low' if the guidelines group have minimum or no dependency on the roles. These guidelines also provide no scientific evaluation and statistical evidence.

Table 2.9 Mapping between recommendation groups and relevance of stakeholder's roles (Source: de Santana et al., 2012)

Guidelines Group	Relevance		
	Developer	**Designer**	**Content Producer**
Navigation	High	High	Medium
Colours	Medium	High	Medium
Text presentation	Medium	High	Medium
Writing	Low	Low	High
Layout	Medium	High	Low
Images and charts	Medium	High	Medium
End user customization	High	Medium	Low
Mark up	High	Low	Medium
Videos and audios	Medium	High	Medium

Guidelines by Rello, Kanvinde, et al. (2012b) was proposed based on eye gaze data. Twenty-two participants with dyslexia and twenty-two participants without dyslexia took part in this experiment. All participants were aged 13 to 37 years old. Questionnaires were used to collect demographic and disability information before a reading task while user preferences were collected after reading task. In the reading task, each participant read two stories that divided into 36 parts within 20 slides. Stories were presented in verse or prose form. Each part has 22 words and is formatted based on parameters value. After that, interview used to collect participants' difficulties encountered in daily reading using different devices. In this experiment, eight parameters with 36 values were used. Results for participants with dyslexia in this experiment with all parameters and its values can be seen in Table 2.10. Bold values were used to represent shortest fixation duration or most preferred by user.

Table 2.10 Results for each parameters value (Source: Rello et al., 2012b)

Parameters	**Value**	**Avg. Fixations Duration (s)**	**(%)**	**User Choice (%)**
Font size	26pt	**0.209**	-	**63.64**
	22pt	0.217	3.8	36.36
	18pt	0.239	14.4	-
	14pt	0.288	37.8	-
Paragraph spacing	3 lines	0.230	4.5	-
	2 lines	**0.220**	-	**63.64**
	1 line	0.242	10.0	-
	0.5 line	0.240	9.1	36.36
Line spacing	1.4 lines	**0.228**	-	**38.64**
	1.2 lines	0.245	7.5	22.73
	1 line	0.240	5.3	34.09
	0.8 line	0.238	4.4	4.55
Character spacing	+14%	**0.205**	-	9.09
	+7%	0.219	6.8	36.36
	0%	0.233	13.7	**38.64**
	-7%	0.233	13.7	15.91
Line length	88 chars per line	**0.215**	-	27.27
	66 chars per line	0.225	4.7	**31.82**
	44 chars per line	0.221	2.8	**31.82**
	22 chars per line	0.230	7.0	9.09
Text grey scale (on white background)	0% (pure black)	0.249	6.0	**72.73**
	25%	0.237	0.9	22.73
	50%	**0.235**	-	4.55
	75%	0.243	3.4	-

Parameters	Value	Avg. Fixations Duration (s)	(%)	User Choice (%)
Background grey scale (on white text)	100% (pure black)	0.255	4.5	**65.91**
	75%	**0.244**	-	15.91
	50%	**0.244**	-	18.18
	25%	0.300	23.0	-
Colour pairs (text/ background)	Black/Cream	**0.214**	-	18.18
	Blue/Yellow	0.220	2.8	6.05
	Dark Brown/Light Mucky Green	0.222	3.7	1.50
	Brown/Dark Mucky Green	0.226	5.6	4.55
	Black/White	0.229	7.0	13.64
	Off-black/Off-white	0.234	9.3	-
	Blue/White	0.238	11.2	18.18
	Black/Yellow	0.239	11.7	**37.86**

Note. Bold value indicates shortest fixation duration or most preferred by participants.

Based on the results on Table 2.10, guidelines for web text was proposed as in Table 2.11. Value for parameters of font size, paragraph and line spacing were proposed because they scored both fixation duration and user choice. However, average values were chosen between two best values for parameters grey scale (both text and background), character spacing and line length because differences in fixation duration and user choice. For colour pairs, Black/Cream was chosen because extreme differences between fixation duration and user choice were recorded. Authors claimed eye gaze data was relevant since user choice might influence by many aspects and might be change in the future.

Table 2.11 Guidelines for the Web (Source: Rello et al., 2012b)

Parameters	Value
Font size	26pt
Paragraph spacing	2.0
Line spacing	1.4
Character spacing	+7%
Text grey scale	10%
Background grey scale	90%
Colour pairs	Black/Cream
Line length	77 characters per line

Even though these guidelines provide scientific evidence, this experiment did not provide any statistical analysis for participants without dyslexia and also no evidence on this guidelines' efficiency. In the design, type of material ('verse' and 'prose') may have an impact on the results and slides were used instead of web pages. According to Zeman, Milton, Smith, & Rylance (2013), different part of brain were used in reading prose and verse. Furthermore, the authors did not mention what type of slides was used. For colour pair parameter, value for each pair in terms of RGB or HEX was not specified. By referring to colour name, it is vague for developer to implement this guideline in their web sites.

It is also unclear which font sizes, spacing were used in relation to the rest of other parameters. In addition to this, parameters were tested independently on different or same slides. For the purpose of combining all parameters values as specified in Table 2.11, I had developed a web page by implementing the guidelines on 23.6-inch monitor as in Figure 2.31. There is no direct way to implement paragraph spacing between `<p>` html tag in CSS, therefore the figure used default spacing between `<p>` tag. Alternatively, `margin-top` or `margin-bottom` can be added between `<p>` tag, using unit cm, mm, in, px, pt and pc. Furthermore, it is unclear which value with which unit is equivalent to 2.0 paragraph spacing in the slide. Since this guideline is focusing on web, it is recommended to use relative font size such as em and percentage instead of absolute font size to display on computer screen. Not only that, both grey scale for text and background were contradict to colour pair recommended as 10% text grey scale were tested on white background, while 90% background grey scale were tested with white text. It is unclear whether 10% text grey scale will have same benefit when combined with Cream background and Black text with 90% background grey scale.

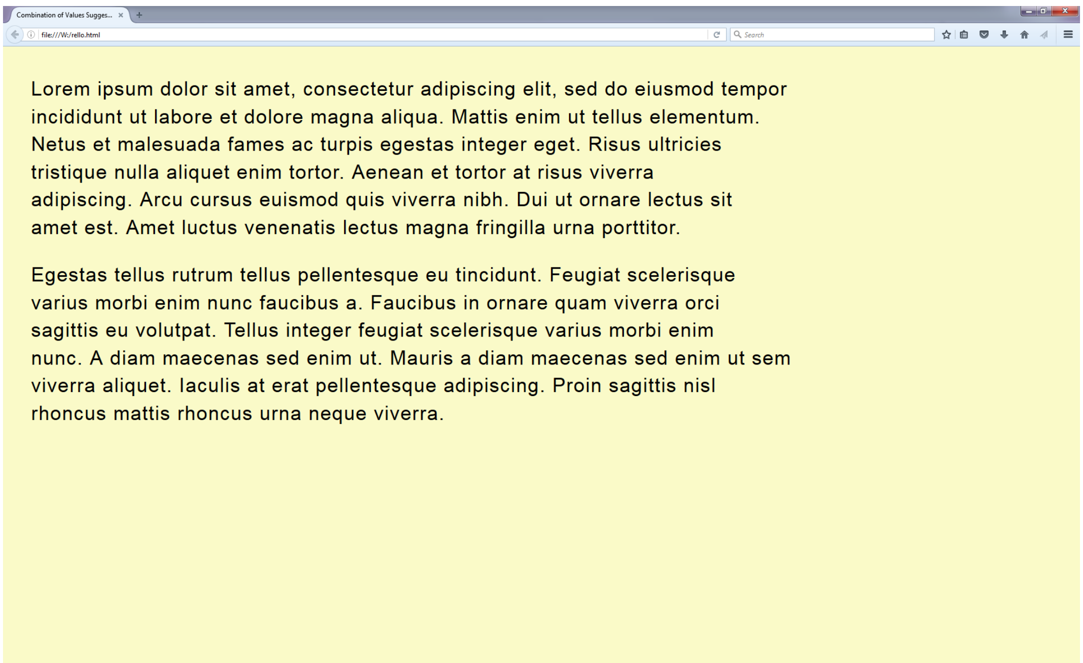


Figure 2.31 Combination of all values in Rello, Kanvinde, et al.’s (2012b) guidelines

Later another set of guidelines was proposed which had more participants, counterbalanced tasks and statistical analysis for all participants (Rello & Baeza-Yates, 2015). 46 participants with dyslexia (aged 11 – 45) and 46 participants without dyslexia (aged 13 – 37) were involved. This study involved similar eight parameters and similar thirty-six conditions, as in previous guidelines (Rello et al., 2012b). The conditions for all parameters are similar except for Grey scale for background, which now had same percentage values as in Grey scale for text; 0%, 25%, 50% and 75%. All parameters were conducted into eight separate experiments and there was no combination between parameters. Figure 2.32 shows example of some parameters such as text grey scale on white background, white text on background grey scale, colour pairs and character spacing used in this study.

Same texts, materials and procedures as in previous study on guidelines were used (Rello et al., 2012b). Default presentation for the texts were black text on white background, Arial 20pt size, left-justified, 0% character spacing, 1.0 line spacing and paragraph spacing, and 66 displayed characters per line. Fixation duration and preference rating were measured as dependent variables while comprehension score as control variables.

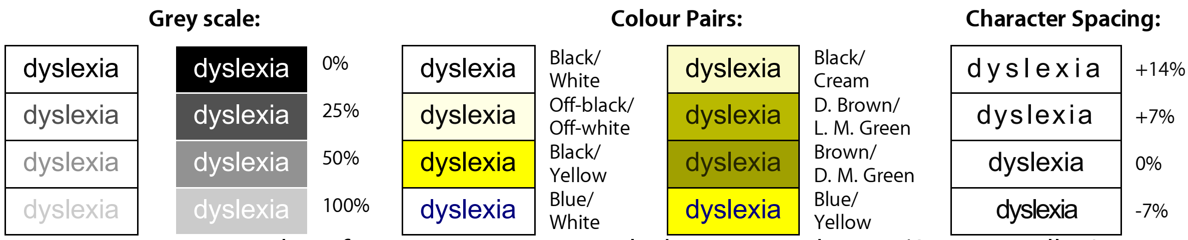


Figure 2.32 Examples of some parameters including its conditions (Source: Rello & Baeza-Yates, 2015)

Results for this study were summarised in Table 2.12. The table shows all significant and non-significant effects for both groups and within groups, including recommendation made. Not only that, recommendation on font category and font type were also provided based on her previous studies (Rello & Baeza-Yates, 2013, 2016) as discussed in Section 2.3.1. It was concluded that text customisation has positive effects on reading however final customisation for some parameters are left to user preferences.

Table 2.12 Results and recommendations for more readable text (Source: Rello & Baeza-Yates, 2015)

Parameters	Measurements	Values with positive effects		Recommendation
		Dyslexic	Non-Dyslexic	
Font size	Fixation duration	Both groups: **18pt, 22pt, 26pt**		Ranging from 18pt to 26pt
		22pt, 26pt	**26pt**	
	Preference rating	**26pt**	**26pt**	
Character spacing	Fixation duration	Both groups: **+7%, +14%**		0%, +7% to +14%
		ns	ns	
	Preference rating	ns	**0%**	
Line length	Fixation duration	ns	ns	-
	Preference rating	ns	**44 chars per line**	
Text Grey scale (white background)	Fixation duration	ns	ns	Black text on White background, or White text on Black background
	Preference rating	**0% (black)**	**0% (black)**	
Background Grey scale (white text)	Fixation duration	ns	ns	
	Preference rating	**0% (black)**	**0% (black)**	
Typeface category	Fixation duration	**Roman** **Sans Serif** **Monospaced**	**Roman** **Sans Serif** **Monospaced**	Roman and Sans Serif
	Preference rating	**Roman** **Sans Serif** ns	**Roman** ns **Proportional**	
Typeface	Fixation duration	**Arial** **Courier** **CMU** **Helvetica**	**Arial** **Courier** **CMU** **Verdana**	Arial, Courier, CMU, Helvetica, Verdana
	Preference rating	**Verdana** **Helvetica** **Arial**	**Verdana** **Helvetica** **Arial**	

Note. ns = no significant effect; bold value = significant effect; CMU = Computer Modern Unicode typeface. Paragraph Spacing, Line Spacing and Colour Pair were not included as they had no significant effect on fixation duration and preference rating for both groups; therefore, no recommendations were made.

In Rello's thesis (2014b), another set of guidelines were proposed based on her previous works (Rello, 2012; Rello & Baeza-Yates, 2012, 2013, 2015, 2016; Rello & Bigham, 2017; Rello et al., 2012b; Rello, Pielot, et al., 2013). The recent guidelines have identical recommendations as in her previous guidelines (2015), except for line spacing where 1.0 to 1.5 were recommended.

It is clear that some limitations in her previous guidelines were overcame in the recent guidelines. However, participants' age in both guidelines had wide gap and no combination of parameters were measured. It also vague whether recommendations made in all of her guidelines will have similar benefits when reading were done on slide and web browser. Not only that, some of the recommendations were contradicted.

All guidelines found in literature review had been summarised in Table 2.13 focusing on text presentation and colours aspects. Furthermore, guidelines discussed in this section had been compiled and mapped into relevant categories as in Table 2.14. In conclusion, some of the recommendations derived from the small corpus of research is confusing and unclear thus create confusion for developers trying to implement websites or other electronic environments suitable for users with dyslexia. Guidelines and recommendations for people with dyslexia are scarce and most of it provides no scientific evidence of its effectiveness for dyslexia in reading. Those guidelines either just produce recommendations or just rework previous recommendations.

Table 2.13 Summary of text presentation and colour pair from various guidelines

Parameters		Zarach (2002)	Friedman & Bryen (2007)	Evett & Brown (2005)	British Dyslexia Association (2018b)	de Santana et al. (2012)	Rello, Kanvinde, et al. (2012b)	Rello (2014b)	Rello & Baeza-Yates (2015)
Colour pair (text/ background)		-	-	Dark Blue/Pale Blue or Black/Yellow	Dark/Light (not white)	Dark/Pastel	Black/Cream	Black/White or White/Black	Black/White or White/Black
Grey scale	Text	-	-	-	-	-	10%		
	Background	-	-	-	-	-	90%		
Typeface category		Sans Serif	Sans Serif	Sans Serif	Sans Serif	Mono-spaced, Sans Serif	-	Mono-spaced, Sans Serif	Sans Serif
Typeface		Arial, Verdana	Arial, Verdana, Helvetica, Tahoma	Arial, Universe, New Century Schoolbook, Comic Sans	Arial, Comic Sans, Verdana, Tahoma, Century Gothic, Trebuchet, Calibri, Open Sans	Arial, Comic Sans, Verdana, Tahoma, Century Gothic, Georgia, Trebuchet	-	Arial, Courier, Computer Modern, Helvetica, Verdana	Arial, Courier, Computer Modern, Helvetica, Verdana
Font size		> 12pt	12pt – 14pt	12pt – 14pt	12pt – 14pt 1-1.2em / 16 – 19px	12pt – 14pt 12px – 14px	26pt	18pt – 22pt	18pt – 26pt
Spacing	Character	-	-	Wider	35%	-	7%	7% to 14%	0%, 7%, 14%
	Line	-	-	1.5 – 2.0	1.5	1.5 – 2.0	1.4	1.0 – 1.5	-
	Paragraph	-	-	Wider	-	-	2.0	-	-
Line length		-	-	60 – 70cpl	60 – 70cpl	60 – 70cpl	77cpl	-	-

Table 2.14 Compilation of categories covered in guidelines

Guidelines	Navigation	Colours	Text Presentation	Heading and Emphasis	Writing Style	Layout	Images, Chart, Videos and Audios	Customisation	Screen Reader
Zarach (2002)	*		*		*	*	*	/	/
Friedman & Bryen (2007)	/	*	/	*	*	*	*	*	*
Evett & Brown (2005)	*	/	/	/	/	/	/	/	/
British Dyslexia Association (2018b)		/	/	/	/	/	*	/	
de Santana et al. (2012)	/	/	/	/	/	/	/	/	/
Rello, Kanvinde, et al. (2012b)		/	/						
Rello (2014b)		/	/						
Rello & Baeza-Yates (2015)		/	/						

Note. '/' indicates categories were mostly covered in the guidelines; '' indicates categories were slightly or partially covered in the guidelines.*

2.6 Current Assistive Technologies Created for People with Dyslexia

Nowadays, people are relying heavily to computers in completing daily basic task such as communication, writing and reading. As people with dyslexia having difficulty in reading and writing, it is important to provide technologies that can accommodate their needs. Lots of assistive technologies had been created for people with dyslexia such as spell checker, text-to-speech software and electronic reader applications.

Spell checker is a program, which recognises words that have been typed and will compares to its dictionary. If the words were not in the dictionary, the program will provide suggestions that closest to the word typed. Some of spell checker programs are capable of grammar checking and punctuation such as MS Office Word. MS Office Word is a document-editing program that integrates spell checker capabilities in a real time. Examples of spell checker programs are Oribi Verity Spell, Ginger, Ghotit, and Franklin Spellchecker (BDA Technology, 2015a).

Text-to-speech software is a program that synthesise text using generated voice. It helps readers with dyslexia to read faster and also helps in reducing their visual stress (BDA Technology, 2015b). This program comes with screen reader software and document-editing programs or it also can be installed stand-alone. Screen reader is a program that attempt to convey interface that currently being displayed to user and will deliver the interface through text synthesiser or text-to-speech program. Examples of text-to-speech programs are ReadingPenTS Oxford, DiTres, Claro Read Standard, Balabolka, SpeakIt, and Write:OutLoud (BDA Technology, 2015b; Rello et al., 2012b).

Electronic readers are device that was designed to read document and can be use with specific applications developed primarily for people with dyslexia. This technology display documents by applying guidelines for dyslexia and provide customisation capabilities for readers. Some people with dyslexia were reported to have better reading speed and comprehension when reading with small handheld e-readers compared with paper (Schneps, Thomson, Chen, Sonnert, & Pomplun, 2013). Furthermore, readers without dyslexia can also benefit from this application as well. Examples of applications are IDEAL, Text4All, and DysWebxia (Kanvinde, Rello, & Baeza-Yates, 2012; Rello, Baeza-Yates, Saggion, Bayarri, & Barbosa, 2013; Rello, Bayarri, et al., 2013; Rello et al., 2012a). Lukeš (2015) had recommended several features for an ideal reader for people with dyslexia, (1) text-to-speech interface with navigation and voice controls, (2) accessible text customisation with sensible initial setting, (3) outline-based navigation, (4) bookmarking, highlighting and annotation functions, (5) chunking interface to display text in smaller segments, including auto scroll, (6) additional advance reader settings, (7) library and access to online repositories of texts.

2.7 Difficulties Encountered by People with Dyslexia When Using Web Sites

In an investigation done by Disability Rights Commission or DRC (2004) with 100 web sites revealed 5 key problems with experienced by people with dyslexia, (1) unclear and confusing web page layout, (2) unclear and confusing navigation mechanisms, (3) inappropriate use of colours, (4) small graphics and text, and (5) complicated language. This investigation used automatic evaluation tools and in-depth evaluation with experts. In explaining this section, problems found in literature reviews were categorised based on key problems found by DRC and will be discussed in the next five sub-sections.

2.7.1 Unclear and Confusing Web Page Layout

Even though this problem is highlighted by DRC as one of key problems, they did not mention in which aspect of layout that is unclear and confusing for dyslexic people. However, author found no literature that discuss and measure how this problem can affect people with dyslexia. Therefore, author considered this as one of the gaps found in the research between dyslexia and accessibility.

2.7.2 Unclear and Confusing Navigation Mechanisms

In a study done by Al-Wabil et al. (2007) as discussed in Section 2.4, it focused on navigation behaviours of people with dyslexia. People with dyslexia had difficulties with scrolling and they preferred to have a navigation bar that is visible at all times. If possible, layout and content should be presented without the needs of scroll bar as people with dyslexia felt it cluttered their screen (Dickinson, Gregor, & Dickinson, 2003). They also had bad experiences with dynamic navigation bar, where it only appeared when hovered by mouse. People with dyslexia had difficulties to control their mouse movement especially when they also had dyspraxia and this was supported by Ismail & Jaafar (2011). However, some readers with dyslexia felt comfortable with the idea of dynamic menus, where information was presented in smaller chunks thus reducing visual complexity (Al-Wabil et al., 2007). Semi-transparent navigation bar makes people with dyslexia had difficulties to read the menu because text on the background exacerbated their reading.

Al-Wabil et al. (2007) discovered that most people with dyslexia are familiar with breadcrumb trails in navigation. However, some of them found it not useful in understanding their location in the web sites. They preferred to click 'Back' browser button to visually see previous web pages because of they did not carry visual image of what they saw in previous page. In addition to this, people with dyslexia preferred to use 'Back/Forward' browser buttons and avoid local 'Back/Forward' buttons.

Site map often provide comprehensive web structure but almost half of people with dyslexia did not noticed and familiar with it. Additionally, some of them felt it worsen their visual disorientation. Even site index serves the same purpose as site map, most of readers with dyslexia found it helpful in information seeking. However, some of them had difficulties in locating the information they were looking for because of their spelling problems and alphabet sequencing problems (Al-Wabil et al., 2007).

2.7.3 Inappropriate Use of Colours

Inappropriate use of colours can bring visual stress and distortions to people with dyslexia. Visual stress is an uncomfortable condition during reading and associated with Meares-Irlen Syndrome. Meares-Irlen is a problem with visual processing and lead to glares, headache, letters movement, blurring and flickering during reading. Even though this syndrome in not a character of people with dyslexia, 50% of reading problems of people with dyslexia is caused by Meares-Irlen syndrome (Department for Education and Skills, 2004; Dickinson et al., 2003; Dyslexia ScotWest, n.d.; Irlen UK, n.d.; McCarthy & Swierenga, 2010; Pijpker, 2013). According to Gregor & Dickinson (2007), this problem can be managed with coloured lenses or filters.

Guidelines and recommendations available scarcely provided statistical evidence on ideal font colour and background colour for people with dyslexia. In a study of development and evaluation of various prototypes of SeeWord, people with dyslexia chose variety colour combinations in order to suit their reading needs (Dickinson, Gregor, & Newell, 2002; Gregor et al., 2003; Gregor & Newell, 2000). Evaluation from this development recorded that colour combinations alleviated participants reading difficulty and reading accuracy was improved compared with default colour when in used. In addition, this study showed no statistic significant between reading with participants' preferred setting and reading errors. Authors claimed that result was caused by individual preferences and they believed manipulation default setting does improve reading (Gregor & Newell, 2000).

In addition to this, most guidelines suggested enable customisation with font and background colours for people with dyslexia (British Dyslexia Association, 2012, 2018b; de Santana et al., 2012; Evett & Brown, 2005; Zarach, 2002). Meanwhile, Checkpoint 2.2 in WCAG 1.0 suggested to brightness differences less than 125 and colour differences less than 500 in order to avoid problems with readers with colour visibility (Rello & Baeza-Yates, 2012; W3C, 2000). According to Gregor & Newell (2000) and Gregor & Dickinson (2007), people with dyslexia preferred colour combinations with lower contrast. By referring to WCAG 2.0, people with cognitive disabilities requires a minimum colour contrast ratio of 4.5:1 and contrast ratio of 3:1 for font size larger than 18pt (W3C, 2014).

People with dyslexia preferred media with soft pastel colour background such as cream and dark colour for text. White background should not be use because of its brightness, and alternative for this colour is light grey with hexadecimal code of #FFFFE5 (British Dyslexia Association, 2018b; de Santana et al., 2012; Evett & Brown, 2005). Example colour pairs given by de Santana et al. (2012) is dark blue text on beige background. On the other hand, guidelines by Rello, Kanvinde, et al. (2012b) suggested cream background with black text based on eye tracking data and user preferences. Both colour pairs were suggested without any RGB or HEX code. In her recent guidelines, the HEX code for colour pairs were provided (Rello & Baeza-Yates, 2015). Black text on White background or White text on Black background was recommended. However, her recommendation on using white background was contradict with white background restriction in de Santana et al. (2012) and British Dyslexia Association (2018b).

2.7.4 Small Graphics and Text

Instructions and breadcrumb trails with smaller font size are hard to read and understood by people with dyslexia. Not only that, their visual discomfort compromise reading process while unnoticeable breadcrumb trails were not serve its purpose to keep track your location in web sites (Al-Wabil et al., 2007; Ismail & Jaafar, 2011). With smaller graphics or visual aids, people with dyslexia had difficulties to scanning web pages in order to find specific content they need. Difficulties of scanning web page and inadequate text presentation accumulated to 16.6% of problems of users with dyslexia (Freire, Petrie, & Power, 2011).

2.7.5 Complicated Language

According to W3Techs (2015), 54.7% of websites use English as their content language while the rest of the percentage are covered by various languages. There is no guarantee that content provided in a particular website are easy to read or using suitable readability score. In a study of analysing complex English words structures by measuring comprehension, it was concluded that higher education students with dyslexia have significant poor reading comprehension in inferential questions compared to students without dyslexia. However no significance differences between students with and without dyslexia for lateral questions (Simmons & Singleton, 2000).

In an eye-tracking study of text simplification using verbal paraphrase measuring text readability, comprehensibility and perception of people with or without dyslexia (Rello, Baeza-Yates, & Saggion, 2013). This study compared between Spanish texts with lexical verbs (e.g. of text '*to hug*') and lexical verbs plus noun collocation (e.g. '*to give a hug*') in their design. Participants completed a set of inferential questions based on the text. There was no significant effect of 'verb' on readability and comprehensibility of participants with and without dyslexia. People with dyslexia preferred text with simple lexical however they did not read any faster with the simple text.

2.8 Conclusion

The literature focusing on digital text presentation and navigation to support people with dyslexia has shown that scarce research has been done with some of the studies did not provide any statistical significance test. In addition, some of the research derived from small corpus had methodological problems in the design of the experiment. Table 2.15 presents a systematic literature review on papers discussed in Section 2.3 which reviewed the effects of text presentations on reading behaviour of people with and without dyslexia. It can be seen that most of the research focused on languages other than English and some of the research had wide age gap of participants. Little works has focused on reading with English and adult participants with dyslexia.

Not only reading in web sites is important, the process of locating the information in the web pages are crucial in determine user experiences. As discussed in Section 2.4, poor navigation structures are burden to people with dyslexia. However, only two research were found that focus on how people with dyslexia doing web navigation. In addition, only small numbers of participants were involved. Therefore, more research should be done in order to explore and have better understanding on that area.

As discussed in Section 2.5, most of the guidelines were recommended without any statistical test and no studies were done in order to validate the existing guidelines. Even though the recent guidelines were empirically tested, some the recommendation values were unclear to be implemented in web environment. For example, Yellow were used with variety of HEX code. Details of the HEX code can be referred to section 'Colour' in Table 2.15.

Thus, for the remaining studies for this programme of research, I will explore the effects of text presentation when reading and navigation with different types of menu organisation and visibility. For text presentation, I will focus on typeface, font size, line length and line spacing. In general, this programme of research will focus on adult participants with dyslexia on both aspects of text presentation and navigation with English contents. Since eye movements are also influenced by text presentation, eye tracking devices will be used in this research to get insight of participants' behaviour. By using eye tracking, unbiased, objective and quantifiable data can be obtained. Not only that, it also offers visual representation on how people have interacted with the presented stimuli in the research.

Table 2.15 Systematic literature review matrix

Typeface *Authors (Year)*	*Participants*	*Languages*	*Type of Tasks*	*Conditions*	*Measurements*
de Leeuw (2010)	21 adult dyslexics 22 adult non-dyslexics	Dutch	Word reading Non-word reading ** read out loud*	Font • Arial 14pt • Dyslexie 12pt	Reading time (raw score on time) Reading accuracy (number of errors and type) Attitude
Pijpker (2013)	22 children dyslexics • 13 low reading level • 9 high reading level 42 children non-dyslexics • 12 low reading level • 30 high reading level	Dutch	Text reading ** read out loud*	Font • Arial 10pt • Dyslexie 9pt Background colour • Yellow A4 paper • White A4 paper	Reading time (raw score on time) Reading accuracy (number of errors and type)
Kuster et al. (2017)	170 children dyslexics	Dutch	1. Sentence reading 2. Word reading a. Simple syllable b. Complex syllable c. Multisyllabic ** read out loud*	1. Font • Arial 13pt • Dyslexie 12pt 2. Font • Arial 16pt, 14pt • Dyslexie 13pt, 11pt • Times New Roman 16pt, 14pt ** all Dyslexie fonts had reduce spacing*	Reading time (raw score on time) Reading accuracy (number of errors) Preferences
Wery & Diliberto (2017)	12 children dyslexics	English	Letter naming Word naming Non-word reading ** read out loud*	Font • OpenDyslexic 10pt • Arial 12pt • Times New Roman 12pt	Reading speed (letters/minute) Reading accuracy (correct letters/total attempt)
Rello & Baeza-Yates (2013, 2016)	48 mixed children and adult dyslexics 49 mixed children and adult non-dyslexics	Spanish	Text reading ** silent, eye tracking*	Font • Arial 14pt • Arial Italic 14pt • Computer Modern 14pt • Courier 14pt • Garamond 14pt • Helvetica 14pt • Myriad 14pt	Reading time (raw score on time) Number of fixations Average fixation duration Preferences

				• OpenDyslexic 14pt • OpenDyslexic Italic 14pt • Times 14pt • Times Italic 14pt • Verdana 14pt	

Font size					
Authors (Year)	*Participants*	*Languages*	*Type of Tasks*	*Conditions*	*Measurements*
O'Brien et al. (2005)	22 children dyslexics 12 children non-dyslexics	English	Sentence reading ** read out loud*	Size • -0.2 to 1.0 Critical Print Size	Reading speed (words/minute)
Rello, Pielot, et al. (2013)	24 mixed children and adult dyslexics	Spanish	Text reading ** silent, eye tracking*	Size • 10pt, 12pt, 14pt, 18pt, 22pt, 26pt Line spacing • 0.8, 1.0, 1.4, 1.8	Average fixation duration Comprehension score Perception
Rello, Kanvinde, et al. (2012b)	22 mixed children and adult dyslexics 22 mixed children and adult non-dyslexics	Spanish	Text reading ** silent, eye tracking*	Size • 14pt, 18pt, 22pt, 26pt	Average fixation duration Preferences
Rello & Baeza-Yates (2015)	46 mixed children and adult dyslexics 46 mixed children and adult non-dyslexics	Spanish	Text reading ** silent, eye tracking*	Size • 14pt, 18pt, 22pt, 26pt	Average fixation duration Preferences
Schoonewelle (2013)	39 high school dyslexics	Dutch	Text reading ** silent*	Font Size • 8.5pt, 10pt Letter Spacing • 0pt, 0.5pt Line Spacing • 1.0, 1.15	Comprehension score Preferences

Line length					
Authors (Year)	*Participants*	*Languages*	*Type of Tasks*	*Conditions*	*Measurements*
Schneps, Thomson, Sonnert, et al. (2013)	27 high school dyslexics	English	Text reading ** silent, eye tracking*	Hand placement • Held/not held Character spacing • Normal/Extra wide (29%) Device	Reading rate (words/minute) Number of fixations Number of inefficient saccades Fidelity score

Authors (Year)	Participants	Languages	Type of Tasks	Conditions	Measurements
				• iPod (12.7cpl) / iPad (67.2cpl) * *iPod had shorter line length*	
Rello, Kanvinde, et al. (2012b)	22 mixed children and adult dyslexics 22 mixed children and adult non-dyslexics	Spanish	Text reading * *silent, eye tracking*	Line length • 22, 44, 66, 88 characters per line	Average fixation duration Preferences
Rello & Baeza-Yates (2015)	46 mixed children and adult dyslexics 46 mixed children and adult non-dyslexics	Spanish	Text reading * *silent, eye tracking*	Line length • 22, 44, 66, 88 characters per line	Average fixation duration Preferences
Line spacing and other spacing					
Authors (Year)	*Participants*	*Languages*	*Type of Tasks*	*Conditions*	*Measurements*
Rello, Pielot, et al. (2013)	24 mixed children and adult dyslexics	Spanish	Text reading * *silent, eye tracking*	Size • 10pt, 12pt, 14pt, 18pt, 22pt, 26pt Line spacing • 0.8, 1.0, 1.4, 1.8	Average fixation duration Comprehension score Perception
Rello, Kanvinde, et al. (2012b)	22 mixed children and adult dyslexics 22 mixed children and adult non-dyslexics	Spanish	Text reading * *silent, eye tracking*	Line spacing • 0.8, 1.0, 1.4, 1.8	Average fixation duration Preferences
Rello & Baeza-Yates (2015)	46 mixed children and adult dyslexics 46 mixed children and adult non-dyslexics	Spanish	Text reading * *silent, eye tracking*	Line spacing • 0.8, 1.0, 1.4, 1.8	Average fixation duration Preferences
Schoonewelle (2013)	39 high school dyslexics	Dutch	Text reading * *silent*	Font Size • 8.5pt, 10pt Letter Spacing • 0pt, 0.5pt Line Spacing • 1.0, 1.15	Comprehension score Preferences
Zorzi et al. (2012)	1. 34 children dyslexics (Italian) 40 children dyslexics (French)	Italian French	1. Text reading 2. Text reading * *read out loud*	1. Letter spacing • 2.7pt with line spacing 1.0 • 5.2pt with line spacing 2.0 and word spacing 3x	Reading rate (syllables/second) Reading accuracy (number of errors)

	30 children non-dyslexics 2. 20 children dyslexics (Italian)			** bigger line spacing to balance text appearance* 2. Letter spacing • 2.7pt with line spacing 2.0 • 5.2pt with line spacing 2.0 and word spacing 3x	
Schneps, Thomson, Sonnert, et al. (2013)	27 high school dyslexics	English	Text reading ** silent, eye tracking*	Hand placement • Held/not held Letter spacing • Normal/Extra wide (29%) Device • iPod (12.7cpl) / iPad (67.2cpl)	Reading rate (words/minute) Number of fixations Number of inefficient saccades Comprehension score
Spinelli et al. (2002)	27 children dyslexics 99 children non-dyslexics	Italian	Letter naming ** read out loud, eye tracking*	Letter spacing • 0.24°, 0.32°, 0.41°, 0.59° ** spacing was manipulated via centre-to-centre between letters* Blurring effect • Default, 2-fold blur, 4-fold blur ** blurred using de-focus tool in Adobe Photoshop program*	Reading speed
Martelli et al. (2009)	29 children dyslexics 32 children non-dyslexics	Italian	Letter identification ** silent, eye tracking*	Letter spacing • Increment of centre-to-centre from target letter	Target accuracy
Perea et al. (2012)	20 children dyslexics 20 children non-dyslexics	Spanish	Text reading ** read out loud*	Letter spacing • Default, +1.2pt	Reading rate (words/minute) Reading accuracy Comprehension score
Moll & Jones (2013)	17 adult dyslexics 17 adult non-dyslexics	English	Letter naming ** read out loud, eye tracking*	Letter spacing • 1.0°, 2.5°, 5.0° visual angle ** spacing was manipulated via distance between letters*	Reading time (gaze duration)
Sjoblom et al. (2016)	24 adult dyslexics 24 adult non-dyslexics	English	Text reading ** read out loud*	Letter spacing • 2.7pt with line spacing 1.0 • 5.2pt with line spacing 2.0 ** bigger line spacing to balance text appearance*	Reading rate (syllables/minute) Reading accuracy (number of errors)

				Colour overlay • Without overlay • 1 out of 10 colour overlays – yellow, orange, magenta, pink, purple, sky blue, aqua blue, grass green, jade green, celery green	
Colour					
Authors (Year)	*Participants*	*Languages*	*Type of Tasks*	*Conditions*	*Measurements*
Rello & Baeza-Yates (2012)	23 mixed children and adult dyslexics 92 mixed children and adult non-dyslexics	Spanish	Text reading ** silent, eye tracking*	Colour (text/background) • Black (#000000)/white (#FFFFFF) • Black (#000000)/yellow (#FFFF00) • Black (#000000)/cream (#FAFAC8) • Off-black (#0A0A0A)/off-white (#FFFFE5) • Blue (#00007D)/white (#FFFFFF) • Dark brown (#1E1E00)/light green (#B9B900) • Brown (282800)/dark green (#A0A000) • Blue (#00007D)/yellow (#FFFF00)	Average fixation duration Preferences
Rello & Bigham (2017)	89 adult dyslexics 252 adult non-dyslexics	Spanish	Text reading ** silent, online environment*	Background colour • Blue (#96ADFC) • Blue grey (#DBE1F1) • Grey (#D8D3D6) • Green (#A8F29A) • Turquoise (#A5F7E1) • Purple (#B987DC) • Red (#E0A6AA) • Peach (#EDD1B0) • Orange (#EDDD6E) • Yellow (#F8FD89)	Reading time (raw score on time) Mouse distance (total pixels that the m0use travelled over text)

Pijpker (2013)	22 children dyslexics • 13 low reading level • 9 high reading level 42 children non-dyslexics • 12 low reading level • 30 high reading level	Dutch	Text reading ** read out loud*	Font • Arial 10pt • Dyslexie 9pt Background colour • Yellow A4 paper • White A4 paper	Reading time (raw score on time) Reading accuracy (number of errors and type)
Rello, Kanvinde, et al. (2012b)	22 mixed children and adult dyslexics 22 mixed children and adult non-dyslexics	Spanish	Text reading ** silent, eye tracking*	Colour (text/background) • Black/cream • Blue/yellow • Dark brown/light mucky green • Brown/dark mucky green • Black/white • Off-black/off-white • Blue/white • Black/yellow ** no HEX/RGB code*	Average fixation duration Preferences
Rello & Baeza-Yates (2015)	46 mixed children and adult dyslexics 46 mixed children and adult non-dyslexics	Spanish	Text reading ** silent, eye tracking*	Colour (text/background) • Black (#000000)/cream (#FAFAC8) • Blue (#00007D)/yellow (#FFFF00) • Dark brown (#1E1E00)/light green (#B9B900) • Brown (#282800)/green (#A0A000) • Black (#000000)/white (#FFFFFF) • Off-black (#A0A0A0)/off-white (#FFFFE5) • Blue (#00007D)/white (#FFFFFF) • Black (#000000)/yellow (#FFFF00)	Average fixation duration Preferences

Note. For Colour Pair, value in the parentheses uses HEX code.

CHAPTER 3: STUDY 1 – EFFECTS OF FONT SIZE AND TYPEFACE ON DETAILED READING

3.1 Introduction

This chapter describes Study 1 in this programme of research which aims to answer Research Questions 1(a) and 1(b) discussed in Chapter 1:

1 (a) To what extent does the typeface used in the presentation of text on a computer screen affect eye gaze behaviour, reading performance, comprehension, preferences and opinions of adults with dyslexia compared to adults without dyslexia?

1 (b) To what extent does the font size used in the presentation of text on a computer screen affect eye gaze behaviour, reading performance, comprehension, preferences and opinions of adults with dyslexia compared to adults without dyslexia?

As discussed in Chapter 1, people with dyslexia show varying symptoms from individual to individual with different levels of severity and different manifestations of dyslexia. In addition to this, they may show inconsistent results in different dyslexia assessments (Rello et al., 2012b; Riddick et al., 1997; Stanovich, 1996). Therefore, it is important to distinguish people with dyslexia into some finer grained categories; in this programme of research, I have divided them in those who have mild, and moderate dyslexia with a well-establish diagnostic tool. This will allow me to investigate whether there are any differences due to severity of dyslexia.

As was seen from the discussion in Chapter 2, a number of studies have investigated which typefaces and font sizes might benefit people with dyslexia (de Leeuw, 2010; Harley et al., 2013; Kuster et al., 2017; Marinus et al., 2016; O'Brien et al., 2005; Pijpker, 2013; Rello & Baeza-Yates, 2013, 2015, 2016; Rello et al., 2012b; Rello, Pielot, et al., 2013; Wery & Diliberto, 2017). However, as discussed, there is no empirical evidence for the benefits of the dyslexia-optimised typefaces for people with dyslexia. Therefore Study 1 is designed to investigate the effect of typefaces (serif, sans serif and dyslexia-optimised) and font sizes on eye gaze behaviour, reading performance, and reading comprehension of people with dyslexia and people without dyslexia during detailed reading. For eye gaze behaviour, fixation duration and number of fixations will be measured. An eye tracking device will be used in order to measure eye gaze behaviour. For reading performance, reading time and comprehension scores will be measured. For preferences, ease of reading, preferred combination of typeface and font size, previous engagement with dyslexia-optimised typefaces, rating of interest and rating of comfort using dyslexia-optimised typefaces will be measured.

As reading comprehension is measured in this study, it is necessary to ensure that all questions asked about the texts read have a similar level of difficulty. A preparatory study was therefore conducted to test the accuracy and difficulty of comprehension questions to be used. This chapter will first present the method and results of the preparatory study. Materials prepared in the preparatory study will be used in the Study 1. The rest of this chapter will also describe the method, results and discussion of Study 1.

3.2 Preparatory Study – Assessment of Questions Difficulty

According to Davey (1988), the validity of comprehension assessment can be affected by at least three groups of variables which are passage features, question types and question formats.

- Passage features refers to word difficulty (such as number of words in a passage, syllable length, frequency of words and types of word) and syntactic difficulty (such as sentence length and number of clauses per sentence)
- Question type classifies questions based on inference type and 'location' of answer (Davey used the term 'location' to refer whether the answer is explicitly or implicitly stated in the passage, not the position of the answer in the text)
- Question format describes multiple-choice versus open ended questions, interrogative word ("wh–" question) versus incomplete sentence, question length, correct answer length, incorrect answer length and so on

Davey (1988) ran a study with poor and normal readers, in which all participants were asked to read 11 texts and answer 72 comprehension questions. All the questions were coded into the three groups of variables discussed above. Regardless of the reading ability of the participants, question length and location of answer contributed to the prediction of question difficulty. Furthermore, the study found that passage feature variable did not contribute significantly to question difficulty. It had been expected that passage feature would play an important role in aiding readers' comprehension.

Fourteen texts were developed to be used in the Study 1 and their characteristics are discussed in Section 3.3.1.3. Considering the findings from Davey (1988), multiple choice questions were developed for each text. The preparatory study aims to test the accuracy and difficulty of the comprehension questions to be used in the Study 1.

Initially, participants were asked to read texts and answer proposed multiple choice questions for each of the texts for Study 1. They needed to answer and rate the questions twice, on the first occasion without referring to the text, to assess the difficulty of the question; and on the second occasion while referring back to the text, to assess the difficulty of understanding the question and locating the answer to the question in the text. This enabled me to make a selection of which questions would be used in Study 1.

3.2.1 Method

3.2.1.1 Design

A repeated measures design was chosen to investigate the difficulties of the questions, how accurately participants could answer the questions and whether they already knew the answer for the questions before reading the texts. Two types of difficulties were measured: Question Difficulty Rating (without referring to the text) and Answer Difficulty Rating (while referring back to the text).

Accuracy was measured by participants' correct answers to the multiple choice questions. Participants' familiarity with the subject matter of the questions was measured by asking whether they already knew the answer before reading the text (the Previous Knowledge Assessment).

As will be discussed in Section 3.2.2, a conflict arose with questions for two texts in the analysis of the preparatory study. Therefore, the preparatory study was repeated for these two texts with comprehension questions modified or replaced. The second round of the preparatory study applied a similar repeated measures design as the first round.

3.2.1.2 Participants

Participants were all native English speakers. In the first round of the preparatory study, four male participants took part. Their ages ranged from 28 to 38 years old. In the second round of the preparatory study, three male participants (one of them had participated in the previous round) and one female participant took part. Their ages ranged from 20 to 55 years old.

3.2.1.3 Materials

Fourteen texts and questions were developed which would be used for two practice tasks and twelve experimental tasks in Study 1. The comprehension questions for each text were four multiple choice questions with three response alternatives. All texts were adapted from Olympic Games Initiative Movement website[5], Wikipedia entries on the Olympic Games[6] and Kamollimsakul (2014). The characteristics of the texts are described in Section 3.3.1.3.

The comprehension questions were developed based on work by Davey (1988), a correct answer for each question needs to be stated explicitly in the text. Wherever possible, one question was developed for each paragraph in the passage. The questions were presented in a random order regardless of which paragraph they referred to. Each question had one correct answer and two alternative distractor choices. Figure 3.1 shows one of the multiple choice questions and answers.

In the first Olympic football competition, which team beat the team from Athens?

a. **Smyrna**
b. **Denmark**
c. **Argentina**

Figure 3.1 Sample multiple choice question and answers

Each of the comprehension questions was accompanied by two difficulty ratings and a familiarity question. Both difficulty ratings were rated on 9-point Likert items. Participants' familiarity with the subject matter of the questions was measured by asking whether they already knew the answer before reading the text using a simple Yes/No

[5] http://www.olympic.org
[6] https://en.wikipedia.org/wiki/Olympic_Games

option. Appendix B shows the 14 sets of texts and a sample of one set of comprehension questions with its accompanying difficulty ratings and familiarity question.

3.2.1.4 Procedure

A file containing the texts and questions was emailed to participants. They were asked to read the passages and answer the accompanying questions. After they finished reading the text, they first answered the questions without looking at the text and rated the difficulty of each question on a 9-point Likert-scale item (the Question Difficulty Rating). They then looked back at the text and rated the difficulty of each question in relation to the text on 9-point Likert-scale item (the Answer Difficulty Rating) and indicated whether they had previous knowledge of the question before reading the text (the Previous Knowledge Assessment). After they had answered the comprehension questions and difficulty assessments, participants returned the file through email or on paper.

In the second round of the preparatory study, texts and modified questions were emailed to participants. The same procedure as in the first round was applied.

3.2.2 Result and Discussion

For each question in the text, participants' accuracy was averaged into Correct Answer, participants' difficulty rating without referring to the text were averaged into Question Difficulty Rating, participants' difficulty rating while referring back to the text were averaged into Answer Difficulty Rating and participants' familiarity on the subject matter were averaged into Previous Knowledge Assessment.

From the preparatory study, a total of three questions was needed for each text for Study 1. A number of factors contributed to the choice of question. For each text, a question was discarded if it was too difficult, too easy or most of participants had previous knowledge about the topic. Question Difficulty Rating were aimed to have a difficulty rating within range 4.00 to 7.75, Answer Difficulty Rating were aimed to have difficulty rating within range 5.00 to 8.75 and Previous Knowledge Assessment within range 0.00 to 0.50. Based on the criteria, texts which have questions outside these range were included in the second round of preparatory study.

Table 3.1 shows the results for all texts in the first round of the preparatory study. Correct Answer was either 0 (incorrect answer) or 1 (correct answer); Question Difficulty Rating and Answer Difficulty Rating ranged from 0 (very difficult) to 8 (not difficult at all); and Previous Knowledge Assessment was either 0 (no) and 1 (yes). For each text, question marked in red indicates it was discarded from being used Study 1. Mean for all questions that have been selected for each text was also shown in the table. Questions for Experimental Texts 1 and 6 was difficult to choose, therefore both texts will be discussed in the second round of preparatory study.

Table 3.1 Results for first round of preparatory study

Text	**Question**	**Correct Answer (*M*)**	**Question Difficulty Rating (*M*)**	**Answer Difficulty Rating (*M*)**	**Previous Knowledge Assessment (*M*)**
Practice 1	1	0.50	4.75	8.25	0.00
	2	1.00	3.50	6.75	0.25
	3	1.00	4.25	7.00	0.00
	4	0.75	4.50	7.25	0.00
	Selected questions (*M*)	0.75	4.50	7.50	0.00
Practice 2	1	0.75	5.67	7.75	0.00
	2	1.00	6.75	7.75	0.00
	3	0.50	5.50	7.75	0.00
	4	0.75	3.50	5.00	0.00
	Selected questions (*M*)	0.75	5.97	7.75	0.00
Experimental 1	1	1.00	7.50	8.25	0.25
	2	1.00	7.50	8.00	0.00
	3	0.50	5.25	7.00	0.00
	4	0.75	4.67	6.75	0.00
	Selected questions (*M*)	*Results discussed in the second round of preparatory study, see Table 3.2*			
Experimental 2	1	1.00	5.00	8.25	0.00
	2	0.50	4.00	7.75	0.00
	3	1.00	8.25	8.50	0.00
	4	1.00	4.25	6.75	0.00
	Selected questions (*M*)	0.83	4.42	7.58	0.00
Experimental 3	1	0.50	5.25	5.75	0.00
	2	0.75	5.75	6.25	0.00
	3	0.50	5.00	5.50	0.00
	4	1.00	5.25	6.00	0.00
	Selected questions (*M*)	0.75	5.42	6.00	0.00

Text	Question	Correct Answer (*M*)	Question Difficulty Rating (*M*)	Answer Difficulty Rating (*M*)	Previous Knowledge Assessment (*M*)
Experimental 4	1	0.75	6.75	7.25	0.00
	2	1.00	6.75	8.00	0.50
	3	0.50	6.75	7.50	0.00
	4	0.50	5.50	6.50	0.00
	Selected questions (*M*)	0.58	6.33	7.08	0.00
Experimental 5	1	1.00	7.25	8.00	0.25
	2	0.75	7.00	7.67	0.50
	3	1.00	8.00	8.25	0.00
	4	1.00	6.00	7.50	0.50
	Selected questions (*M*)	0.92	6.75	7.72	0.42
Experimental 6	1	0.50	5.50	6.75	0.25
	2	1.00	7.25	7.00	0.50
	3	0.25	5.50	5.50	0.00
	4	0.75	4.50	5.25	0.00
	Selected questions (*M*)	*Results discussed in the second round of preparatory study, see Table 3.2*			
Experimental 7	1	1.00	5.00	8.25	0.00
	2	1.00	6.25	8.00	0.00
	3	0.75	5.00	6.75	0.00
	4	0.50	3.75	5.00	0.00
	Selected questions (*M*)	0.92	5.42	7.67	0.00
Experimental 8	1	1.00	7.00	8.33	0.75
	2	1.00	7.75	8.25	0.00
	3	1.00	5.75	8.50	0.00
	4	1.00	5.50	8.00	0.00
	Selected questions (*M*)	1.00	6.33	8.25	0.00
Experimental 9	1	0.75	5.25	7.50	0.00
	2	1.00	6.75	8.25	0.00
	3	0.75	6.00	8.00	0.00
	4	1.00	5.50	4.50	0.00
	Selected questions (*M*)	0.83	6.00	7.92	0.00
Experimental 10	1	1.00	6.25	6.50	0.00
	2	0.75	7.25	8.50	0.00
	3	1.00	7.25	8.00	0.00
	4	0.75	4.75	7.75	0.00
	Selected questions (*M*)	0.92	6.92	7.67	0.00

Text	Question	Correct Answer (*M*)	Question Difficulty Rating (*M*)	Answer Difficulty Rating (*M*)	Previous Knowledge Assessment (*M*)
Experimental 11	1	0.75	6.00	8.25	0.00
	2	0.75	5.00	6.75	0.00
	3	0.75	5.00	6.00	0.00
	4	1.00	6.25	7.75	0.25
	Selected questions (*M*)	0.83	5.75	7.58	0.08
Experimental 12	1	1.00	6.25	7.75	0.00
	2	1.00	7.50	7.25	0.50
	3	1.00	6.75	8.00	0.00
	4	1.00	7.25	8.00	0.00
	Selected questions (*M*)	1.00	6.75	7.92	0.00

Note. Questions marked in red were discarded.

As mentioned previously, the process of choosing the three final questions was more difficult for two texts (Experimental Texts 1 and 6). Therefore, a second round of preparatory study was conducted. For Experimental Text 1, a conflict arose because there were two questions (Question 1 and Question 2) that were rated too easy by participants in both Question and Answer Difficulty Ratings. Questions for Experimental Text 1 were modified for a second round of preparatory study.

For Experimental Text 6, Question 2 was discarded because half of the participants had previous knowledge about the topic. In addition, three participants answered Question 3 incorrectly. Therefore that question, the correct answer and alternative choices were all reworded. Table 3.2 shows the result for second round of the preparatory study with modified questions for Experimental Text 1 and questions that have been selected for both set Experimental Text 1 and Text 6.

From the selected questions in Table 3.1 and Table 3.2, Experimental Texts 8 and 12 (M = 1.00) had the most accurate questions (correct answers from all participants), while Experimental Text 6 (M = 0.50) had the lowest score. Without looking back at the text, Experimental Text 2 (M = 4.42) had the most difficult questions and Experimental Text 10 (M = 6.92) had the easiest rating. While looking back at the text, participants had the most difficulty in finding answer for Experimental Text 6 (M = 5.83) and Experimental Text 8 (M = 8.25) had the easiest rating. Participants were familiar with questions in some experimental texts (M_{Text1} = 0.25, M_{Text5} = 0.42 and $M_{Text6,\ Text11}$ = 0.08).

Table 3.2 Results for second round of preparatory study

Text	Question	Correct Answer (*M*)	Question Difficulty Rating (*M*)	Answer Difficulty Rating (*M*)	Previous Knowledge Assessment (*M*)
Experimental 1	1	0.75	4.25	7.25	0.25
	2	0.50	5.25	6.75	0.75
	3	0.50	5.00	6.75	0.25
	4	0.75	4.75	7.75	0.25
	Selected question (*M*)	0.67	4.67	7.25	0.25
Experimental 6	1	0.50	5.50	6.75	0.25
	2	1.00	7.25	7.00	0.50
	3	0.25	5.50	5.50	0.00
	4	0.75	4.50	5.25	0.00
	Selected question (*M*)	0.50	5.17	5.83	0.08

Note. Question marked in red indicates it was discarded from being used in Study 1. Correct Answer was 0 = incorrect answer, 1 = correct answer; Difficulty Rating Before and Difficulty Rating After ranged from 0 = very difficult, 8 = not difficult at all; Previous Knowledge was 0 = no, 1 = yes.

The preparatory study allowed me to create sets of comprehension questions with similar levels of difficulty, which would not be familiar to participants. Means and standard deviations of all texts for the Correct Answer was 0.80±0.15, Question Difficulty Rating was 5.74±0.84, Answer Difficulty Rating was 7.41±0.69 and Previous Knowledge Assessment was 0.06±0.12. These sets will now be used in the Study 1. It is also worth highlighting that previous studies discussed in Chapter 2 only measured difficulty of text passages read by participants and none of the studies measured difficulty of the questions which were asked after reading to measure reading comprehension (de Leeuw, 2010; Harley et al., 2013; Kuster et al., 2017; Marinus et al., 2016; Pijpker, 2013; Rello & Baeza-Yates, 2013, 2016; Wery & Diliberto, 2017).

3.3 Study 1

One of the challenges in studying digital text presentation is how to measure font size across platforms, browsers, and devices. Measures such as point size, pixels, ems and percentages are frequently used in order to render the size. Point size is the absolute value for font height and it is traditionally used with printed material, which is useful when the physical output size is known (Rabinowitz, 2015). It has fixed-size units and is not scalable, which means it is rendered inconsistently across different computer platforms and does not allow a user to resize the text through their browser's appearance preferences (although this can be achieved with the zoom function). In addition, it is not flexible for developers to make adjustments for inherited styles[7] in HTML tags (Schaeffer, 2008). For screen displays, it is recommended to use relative font sizes such as em and percentage values. Figure 3.2 shows that relative units inherit their size based on their parent tag (`font-size` of `body`), where 1em, 12pt, 16px and 100% have equivalent size when the parent tag is 100% but texts with relative font size differs when the parent tag is 120%. This means the relative font size is scalable, however, texts with unscalable absolute font size can be enlarging by zooming the whole web page (Kessler, 2010; Schaeffer, 2008).

	`body { font-size: 100%; }`	`body { font-size: 120%; }`
`font-size: 1em`	The quick brown fox	The quick brown
`font-size: 12pt`	The quick brown fox	The quick brown fox
`font-size: 16px`	The quick brown fox	The quick brown fox
`font-size: 100%`	The quick brown fox	The quick brown

Figure 3.2 Different size between absolute and relative unit (Source: Schaeffer, 2008)

There are numerous units for relative font size and the number of units will undoubtedly continue to grow in the future. Each declaration of unit and combination of units implemented leads to different results in screen displays (Aderinokun, 2015). Furthermore, typefaces with same point size may have different heights. Figure 3.3 shows that the height of text at 14pt with Times New Roman typeface is smaller than for Georgia at the same point size. With this is mind, I decided to use optical size based on height displayed on monitor screen instead of any font size unit for measuring the presentation of text in this study. This decision was also strengthened by a conclusion in an internal validity study by Schulz (2016): *"When deciding text size, measuring the optical size – preferably on the display where it will be shown – should make the text presentable"* (p. 338).

The quick brown fox jumps over the lazy dog	(Times New Roman, 14pt)
The quick brown fox jumps over the lazy dog	(Georgia, 14pt)

Figure 3.3 Different height for Times New Roman and Georgia

[7] styles that are taken from external style sheet, where the style information was described

3.3.1 Method

3.3.1.1 Design

This study is a 3 × 3 × 2 × 2 mixed-participants design with participant group as the between participants variable and typeface, typeface examples and font size as the repeated measures variables. Participants were divided into three groups (non-dyslexic, mild dyslexic and moderate dyslexic) on the basis of a diagnostic checklist for dyslexia (Snowling et al., 2012). For the typeface variable, three levels were used: serif, sans serif and dyslexia-optimised. For each level, two examples of typeface were taken. For the serif level, Times New Roman and Georgia were used; for the sans serif level, Verdana and Arial were used; and for the dyslexia-optimised level, OpenDyslexic and Lexie Readable. For font sizes, two levels of size were used: small (x-height of 2.5mm) and medium (x-height of 3.3mm).

The dependent variables were grouped into three types of measurements: (1) eye gaze behaviour, (2) reading performance and (3) participants' preferences and opinions.

(1) The eye gaze behaviour measures are: number of fixations and average fixation duration. Fixation is a point where the eye rests for a period of time. Readability improves when people make fewer fixations and shorter fixation duration (Rello & Baeza-Yates, 2015; Schneps, Thomson, Sonnert, et al., 2013).

(2) The reading performance measures are: reading time and comprehension scores. Comprehension scores were only measured to make sure participants read the text.

(3) The participants' preferences and opinions measures are: rating of ease of reading, preferred combination of typeface and font size, previous engagement with dyslexia-optimised typefaces, rating of interest and rating of comfort using dyslexia-optimised typefaces. All ratings were gathered in a post-study questionnaire on 7-point Likert items.

Participants read 12 texts, one with each combination of typeface example and font size. The order of the texts was counterbalanced.

Ethics approval for this study has been granted by Physical Sciences Ethics Committee of the University of York. Participants were ensured about the confidentiality and anonymity of their data.

3.3.1.2 Participants

Participants were asked to complete an online screening diagnostic tool. They were categorised into non-dyslexic group if their diagnostic scores ranged between 0 to 10, mild dyslexic group between 11 to 21 and moderate dyslexic group between 22 to 32. Total of 36 participants took part in this study, 12 participants in each group (non-dyslexic, mild dyslexic and moderate dyslexic). All participants were native English speakers with normal vision or vision correctable with eyewear.

The non-dyslexic group comprised 5 females and 7 males with ages ranging from 18 to 24 years (*M* = 20.33, *SD* = 2.02). Ten were undergraduate students (with three of them were working part time) and two were Master students. Eleven had more than 6 years of Web experience and one had 4 – 6 years of Web experience. Mean of Web expertise is 5.58 (*SD* = 1.24), on a 7-point Likert item.

The mild dyslexic group comprised 8 females and 4 males with ages ranging from 18 to 23 years (*M* = 19.83, *SD* = 1.59). Twelve were undergraduate students, with one of them were working part time. All participants had more than 6 years of Web experience. Mean of Web expertise is 4.92 (*SD* = 1.24), on a 7-point Likert item.

The moderate dyslexic group comprised 6 females and 6 males with ages ranging from 18 to 31 years (*M* = 22.25, *SD* = 4.22). Eight were undergraduate students, one was Masters student, two were PhD students and one was employed. All participants had more than 6 years of Web experience. Mean of Web expertise is 4.50 (*SD* = 0.80), on a 7-point Likert item.

Participants were given a £10 gift voucher from Amazon or Marks & Spencer to thank them for participating in the study.

3.3.1.3 Materials and Equipment

Figure 3.4 shows the dual monitor setup used in this study. Participants used a wireless keyboard, a wireless mouse and a monitor to read texts and answer questions, while the researcher used a laptop. The laptop was connected to the participant's eye tracker (via USB cable) and participant's monitor (via VGA cable). A Bluetooth USB dongle for the participant's wireless mouse and keyboard were also attached to the laptop.

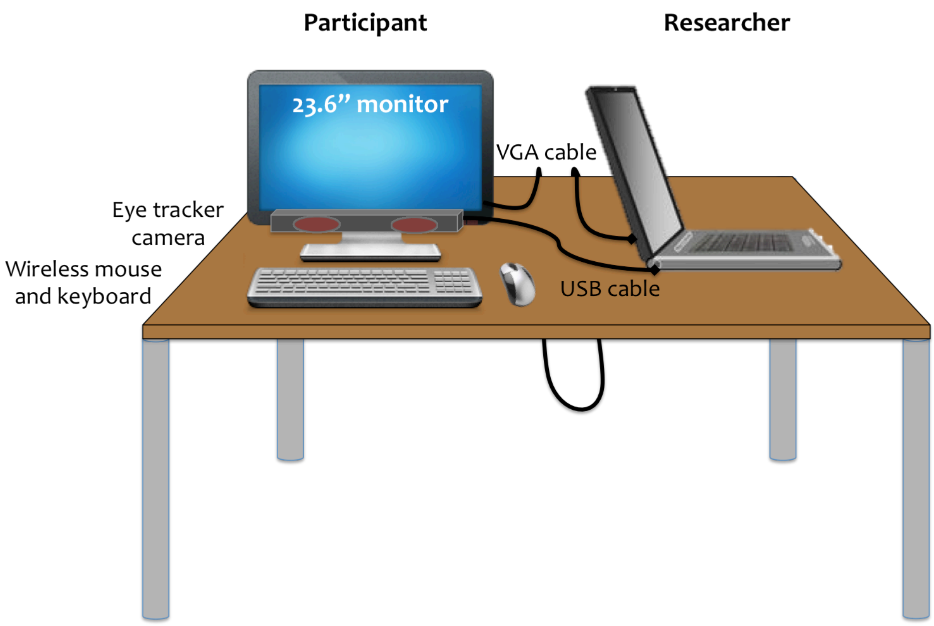


Figure 3.4 Dual monitor setup for Study 1

For the participants, a 23.6-inch monitor screen with screen resolution of 1920px × 1200px and aspect ratio of 16:10 was used to display all materials used in this study. A remote eye tracker camera was attached with a magnetic strip to the monitor screen. The eye tracker, a SMI RED250 Mobile with a sampling rate of 250Hz, was used to capture eye gaze with 5-point calibration.

A viewing distance of approximately 60cm from the screen was maintained. For ecological reasons, a chin rest was not used to maintain an exact viewing distance in this study. Previous studies have shown no difference in reading rate and comprehension score when readers were asked to fixed their head during eye movement studies and when the readers did not (Rayner, 1998).

The participants sat on a fixed chair with seat height adjustment capability. The fixed chair was needed to control the viewing distance from participant and the eye tracker. The height of the chair was also adjusted depending on the participant's height in order to make sure participants' eyes are on the accepted level for the eye tracker.

For the researcher, a DELL Precision M4800 (15.6 inches) laptop running Windows 8.1 Pro on Intel® Core™ i7 with 8GB memory was used to execute the Experiment Center™ program from SMI. The program enabled me to manipulate experiment file and monitor participant's eye gaze.

The online Adult Reading Questionnaire, ARQ (Snowling et al., 2012) was used to screen whether participants were non-dyslexic, mild dyslexic and moderate dyslexic. The questionnaire asked participants' information on their reading habit, literacy, organisation, dyslexia symptoms and its severity. Participants were categorised into non-dyslexic if their scores ranged between 0 to 10, mild dyslexic between 11 to 21 and moderate dyslexic between 22 to 32. The ARQ can be found in Appendix C.

A pre-study questionnaire consisted of questions about participants' use of the web and demographic information. The questions included age, gender, occupation, expertise and the use of the web. The pre-study questionnaire can be found in Appendix E.

Fourteen texts and comprehension questions (consisting of three multiple choice questions for each text) were developed. The questions were developed and tested in the preparatory study described in Section 3.2. Two texts were practice texts and 12 were the experimental texts. In each text, the number of complex words, percentage of complex words and average words/sentence were controlled within ± 10% range from mean value. In order to control these scores between ± 10% range from mean value, texts were repeatedly reworded, replaced and restructured. Texts created were measured with a free readability test tool, Read-Able[8]. Sample of texts can be found in Appendix B.

All texts were adapted from Olympic Games Initiative Movement website[9], Wikipedia entries on the Olympic Games[10] and Kamollimsakul (2014). Each text has the following characteristics:

- Between 270 to 285 words
- Between 14 to 17 sentences
- 4 paragraphs
- 45 to 53 complex words (word with more than three syllables)
- 15.79% to 19.00% complex words
- 16.41 to 19.07 average words per sentence

To ensure all the texts were comparable, a number of readability measures were controlled within ± 10% range from the mean value. The readability measures used comprised the Flesch-Kincaid Grade Level (FKGL), Flesch-Kincaid Reading Ease (FKRE) Test and the Gunning Fog Score (GFS). FKRE is a new formula recalculated from FKGL and provide United States (US) grade level instead of range scores (Colmer, 2019b; Klare,

---

[8] https://www.webpagefx.com/tools/read-able/
[9] http://www.olympic.org
[10] https://en.wikipedia.org/wiki/Olympic_Games

1984). FKRE scores range from 0.0 – 100.0, FKGL score and GFS score equivalent to United States grade level from 0.0 – 20.0 (Armstrong, 1980; Colmer, 2019a, 2019b; Cotugna, Vickery, & Carpenter-Haefele, 2005; Flesch, 1948; Kincaid, Fishburne Jr, Rogers, & Chissom, 1975; Klare, 1984). Readability scores for all the texts and its characteristics are shown in Table 3.3. Scores for the readability measures in the texts are describe below:

- FKRE
  - Score between 47.5 to 54.6, indicates difficult text and represents most general academically oriented articles (Flesch, 1948)
  - Suitable for readers from senior high school level and average readers in first year university
- FKGL
  - Score between 10.1 to 11.1, indicates suitable for US 10th grade to 11th grade and equivalent to UK Year 11 to Year 12 (Cotugna et al., 2005)
  - Easily understood by average university graduates
- GFS
  - Score between 10.6 to 12.9, represents leading magazines articles (Armstrong, 1980)
  - Suitable for senior high school to first year university students

The texts and comprehension questions were presented with six different typefaces in combination with small or medium font size. In order to determine the font size value, 14pt Arial (represents small size) and 18pt Arial (represents medium size) were chosen as a baseline. The character 'x' was printed for both size and the average of x-height was taken 10 times to establish an accurate measure. All texts using small font size had an x-height of 2.5mm and medium font size had an x-height of 3.3mm. Table 3.4 lists the 14 texts with their font combination and a sample of text showing the combination. Practice texts used two other typefaces which were not included in any of the experimental texts.

Table 3.3 Results of readability scores for all sets of texts used in Study 1

Text	No. of Words	No. of Sentences	No. of Complex Words	Complex Words (%)	Average Words/ Sentence	FKRE	FKGL	GFS
P1	281	17	53	18.86	16.53	48.30	10.60	11.00
P2	285	15	52	18.25	19.oo	50.40	10.90	12.70
E1	281	16	48	16.96	17.69	51.10	10.50	11.30
E2	285	15	45	15.79	19.07	51.30	10.80	12.70
E3	278	17	53	19.00	16.41	50.70	10.30	12.30
E4	284	15	46	16.14	19.00	54.00	10.50	11.70
E5	280	16	50	17.86	17.50	49.20	10.70	12.90
E6	283	15	51	18.02	18.37	50.50	10.90	12.10
E7	285	16	45	15.79	17.81	54.60	10.10	11.20
E8	278	16	50	17.99	17.38	47.70	10.90	10.70
E9	285	15	48	16.84	19.00	49.50	11.10	12.70
E10	285	17	45	15.79	16.82	50.00	10.40	10.60
E11	276	15	52	18.77	18.47	48.20	11.10	11.00
E12	275	16	51	18.15	17.56	47.50	11.00	11.70
M	**281.50**	**15.79**	**49.21**	**17.44**	**17.90**	**50.21**	**10.70**	**11.76**
SD	**3.61**	**0.80**	**3.02**	**1.20**	**0.92**	**2.13**	**0.31**	**0.81**

Note. 'FKRE' denotes Flesch-Kincaid Reading Ease; 'FKGL' denotes Flesch-Kincaid Grade Level; 'GFS' denotes Gunning Fog Score; 'P' denotes Practice text; 'E' denotes Experimental text.

Table 3.4 List of all font combinations in practice and experimental texts used for Study 1

Text	Typeface (Example)	Font Size	Sample
Practice 1	Sans serif (Tahoma)	Small	The quick brown fox jumps over the lazy dog.
Practice 2	Serif (Garamond)	Medium	The quick brown fox jumps over the lazy dog.
Experimental 1	Serif (Times New Roman)	Small	The quick brown fox jumps over the lazy dog.
Experimental 2	Sans serif (Verdana)	Medium	The quick brown fox jumps over the lazy dog.

Text	Typeface (Example)	Font Size	Sample
Experimental 3	Dyslexia-optimised (OpenDyslexic)	Small	The quick brown fox jumps over the lazy dog.
Experimental 4	Serif (Georgia)	Small	The quick brown fox jumps over the lazy dog.
Experimental 5	Dyslexia-optimised (Lexie Readable)	Small	The quick brown fox jumps over the lazy dog.
Experimental 6	Serif (Georgia)	Medium	The quick brown fox jumps over the lazy dog.
Experimental 7	Dyslexia-optimised (OpenDyslexic)	Medium	The quick brown fox jumps over the lazy dog.
Experimental 8	Sans serif (Arial)	Small	The quick brown fox jumps over the lazy dog.
Experimental 9	Dyslexia-optimised (Lexie Readable)	Medium	The quick brown fox jumps over the lazy dog.
Experimental 10	Sans serif (Arial)	Medium	The quick brown fox jumps over the lazy dog.
Experimental 11	Serif (Times New Roman)	Medium	The quick brown fox jumps over the lazy dog.
Experimental 12	Sans serif (Verdana)	Small	The quick brown fox jumps over the lazy dog.

Texts were presented in the middle of the participant's screen, left-justified, 80 – 90 characters per line and an average of 8.25mm line spacing for small font or average of 10.42mm line spacing for medium font. They were presented as black text on a white background. Figure 3.5 and Figure 3.6 illustrates one of the texts and the comprehension questions when displayed on the participant's monitor screen. Comprehension questions used the same font size and typeface as the corresponding text.

Two versions of a post-study questionnaire were developed, one for non-dyslexic participants (see Appendix F) and one for dyslexic participants (see Appendix G). Both questionnaires measured participants' ratings of the ease of reading and preferences for the combinations of typefaces and font sizes. In the post-study questionnaire for participants without dyslexia, they were asked about their willingness to receive work from dyslexic colleagues presented using dyslexia-optimised typefaces. In the questionnaire for participants with dyslexia, they were asked in detail about the two dyslexia-optimised typefaces, whether they find them easier to read, whether they would be interested to use them and whether they would be comfortable submitting works using these typefaces. A web page was developed to display excerpts of all the texts they had read using the typeface-size combinations applied to the texts. This companion web page was to assist participants in answering the questionnaire easily. Figure 3.7 shows the companion web page for the questionnaire with all the combinations.

The Ancient Olympic Games

The Ancient Olympic Games were closely linked to religious festivals. They were held every four years at the temple of Zeus in Olympia, Greece. Competition involved representatives of a number of city-states and kingdoms throughout Ancient Greece. It is believed that during the Games, all conflicts among the participating city-states were postponed until the Games were finished. This cessation of hostilities was known as the Olympic peace or truce.

According to historical records, the first Ancient Olympic Games can be traced back to 776 BC. They were dedicated to the Olympian gods and were staged on the ancient plains of Olympia. They continued for nearly 12 centuries, however there is no consensus as to when the Ancient Olympic Games ended. The Ancient Olympic Games gradually declined in importance as the Romans gained power and influence in Greece.

The origin of the Olympics is mysterious and myths identify Heracles and his father Zeus as the founders of the Games. According to legend, it was Heracles who first called the Games "Olympic" and he built the Olympic Stadium in honour of Zeus. Following its completion, he walked in a straight line for 200 steps and called this distance a "stadion", which later became an Ancient Greek and Roman unit of distance.

The Ancient Olympics featured sporting events alongside ritual sacrifices, which honoured Zeus and Pelops, a divine hero and mythical king of Olympia. The winners of events were admired and remembered in poems and statues. The Games featured running events, a pentathlon (consisting of a jumping event, discus throwing, foot races and wrestling), boxing, wrestling, and equestrian events. The Olympic Games were part of a cycle of games known as the Panhellenic Games.

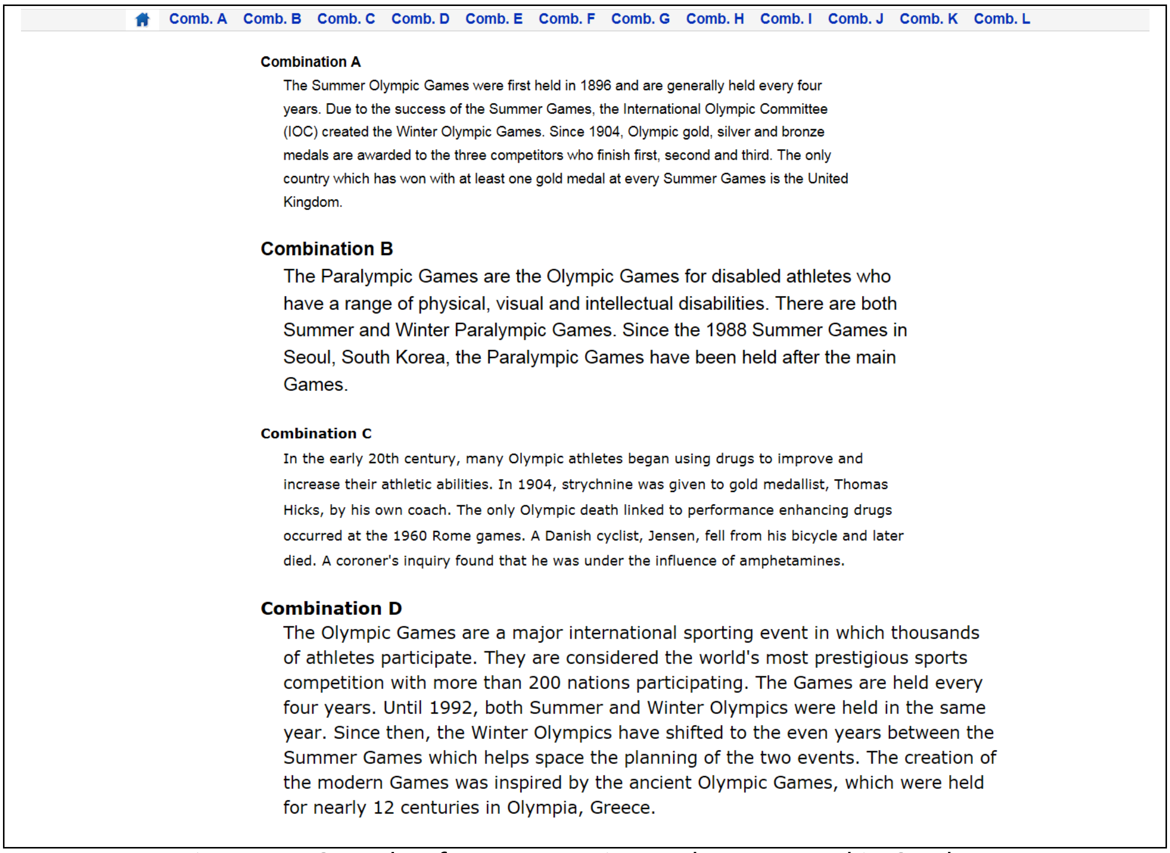
Figure 3.5 Sample of text used in Study 1

Please complete these with 1 answer per question.

1. Pelops was...
   - ○ a son of Zeus
   - ○ an Olympian God
   - ○ the mythical king of Olympia
2. The Ancient Olympics included which kinds of events?
   - ○ Running, pentathlon and equestrian
   - ○ Running, swimming and wrestling
   - ○ Boxing, swimming and equestrian
3. The Ancient Olympics declined because...
   - ○ hostilities amongst the participating city-states
   - ○ the Romans gained power and influence
   - ○ the cult of Heracles and Zeus declined

Next

Figure 3.6 Sample of comprehension questions used in Study 1


Comb. A Comb. B Comb. C Comb. D Comb. E Comb. F Comb. G Comb. H Comb. I Comb. J Comb. K Comb. L

Combination A

The Summer Olympic Games were first held in 1896 and are generally held every four years. Due to the success of the Summer Games, the International Olympic Committee (IOC) created the Winter Olympic Games. Since 1904, Olympic gold, silver and bronze medals are awarded to the three competitors who finish first, second and third. The only country which has won with at least one gold medal at every Summer Games is the United Kingdom.

Combination B

The Paralympic Games are the Olympic Games for disabled athletes who have a range of physical, visual and intellectual disabilities. There are both Summer and Winter Paralympic Games. Since the 1988 Summer Games in Seoul, South Korea, the Paralympic Games have been held after the main Games.

Combination C

In the early 20th century, many Olympic athletes began using drugs to improve and increase their athletic abilities. In 1904, strychnine was given to gold medallist, Thomas Hicks, by his own coach. The only Olympic death linked to performance enhancing drugs occurred at the 1960 Rome games. A Danish cyclist, Jensen, fell from his bicycle and later died. A coroner's inquiry found that he was under the influence of amphetamines.

Combination D

The Olympic Games are a major international sporting event in which thousands of athletes participate. They are considered the world's most prestigious sports competition with more than 200 nations participating. The Games are held every four years. Until 1992, both Summer and Winter Olympics were held in the same year. Since then, the Winter Olympics have shifted to the even years between the Summer Games which helps space the planning of the two events. The creation of the modern Games was inspired by the ancient Olympic Games, which were held for nearly 12 centuries in Olympia, Greece.

Figure 3.7 Sample of accompanying web page used in Study 1

3.3.1.4 Procedure

In order to screen for dyslexia in the participants and assess the severity of their dyslexia, a hyperlink to the dyslexia checklist, the online ARQ (see Appendix C) was emailed to all participants and they completed the questionnaire before coming to the lab to take part in the experiment. The experiment took place in a quiet room where participants were given an explanation of the purpose of the study and the tasks they are going to undertake. They were asked if they had any questions and then completed Section A in the informed consent form (see Appendix D). All participants then completed a demographic questionnaire, covering their use of the Web and personal details (see Appendix E).

Each participant was asked to sit comfortably by adjusting the gap between their body and the table and the height of the fixed chair. The monitor screen was then adjusted so the viewing distance from the participant to the screen was approximately 60cm. The participant was asked to minimise head and body movement during experiment to optimize the eye tracking equipment.

The participant started the experiment by doing the calibration and validation of their eye gaze using the SMI Experiment Center™ application. The calibration was repeated if their eye gaze failed in the validation process. After successful calibration, the application automatically directed participant to the practice texts in a browser window. For practice, the participant read the two practice texts and answered the corresponding set of comprehension questions about the texts. They could also have a short break before starting the experimental texts if they wished.

Calibration and validation of their eye gaze was repeated after the participant took a break, if they had done so, and they had a chance to ask any questions before moving on to the experimental texts. The participant read each experimental text. After reading the text, the participant pressed the spacebar in order to move from the text page to the questions page and answered the questions. They then pressed the spacebar to move on to the next text. The participant read six texts and was then offered a short break before reading the next six texts. After the break, calibration and validation of eye gaze was repeated. The participant was given a chance to ask any questions before each calibration and validation process. Appendix H shows the sequence of texts which participants read.

After reading all the experimental texts, the participant completed the post-study questionnaire (for participants without dyslexia, see Appendix F; for participants with dyslexia, see Appendix G). The participant was shown the companion web page (see Figure 3.7) which presented all the combinations of typefaces and font sizes in order to answer the post-study questionnaire. Participants were then debriefed and any questions they had were answered. Participants were asked to sign Section B in the Consent Form (see Appendix D) to indicate that they were satisfied with all the answers given. Participants were given their gift voucher and asked whether they would like to be sent an example of their eye gaze pattern.

3.3.2 Data Preparation

Two SMI programs were used in order to prepare the data for analyses, SMI Experiment Center™ and SMI BeGaze™. Eye gaze data recorded with SMI Experiment Center™ were exported to SMI BeGaze™. SMI BeGaze™ is used to extract all the data by structuring the data based on participant group and task in the experiment. In addition, SMI BeGaze™ is used to visualise eye gaze patterns.

Using SMI BeGaze™, Area of Interest (AOI) were then created for each paragraph in each text. Results for number of fixations, fixation duration and text reading time were taken from data recorded with SMI Experiment Center™. Comprehension scores were taken from a MySQL database which was recorded during participants' interaction with the web pages. Their preferences were taken from the post-study questionnaire given after the experimental task.

The number of fixations, fixation duration and reading time was calculated from the time when the participant's eye gaze entered their first AOI in the text to the time when the participant pressed the spacebar to move to the comprehension questions web page. Fixations outside the AOIs were not counted. Average fixation duration was calculated by dividing the fixation duration and number of fixations in AOIs.

The results were grouped into three categories: (1) eye gaze behaviour (number of fixations and average fixation duration), (2) reading performance (text reading time and comprehension score) and (3) participants' preferences and opinions.

Each text read by participants was visually inspected using SMI BeGaze™ in order to confirm that the eye tracker had correctly recorded their eye gaze. If their eye gaze was not accurate, due to excessive head movement or other failures, their data was excluded from the analysis. The inclusion criteria for each participant's eye gaze data in this study are: (1) eye gaze was accurately recorded with at least two paragraphs per text and, (2) eye gaze was accurately recorded with at least ten texts out of all twelve texts. Based on these criteria, six initial participants were replaced with new participants for whom age, gender and their severity of dyslexia were matched.

Since the number of words in each paragraph were not the same across all the texts, all data recorded using SMI Experiment Center™ were divided based on the number of words within the paragraphs for which a participant had accurately recorded data. That is to say, the number of fixations, the average of fixation duration and reading time were measured per word. In addition, texts under the same category of typeface were averaged. Texts read with small Times New Roman and small Georgia were grouped as 'serif-small', medium Times New Roman and medium Georgia were grouped as 'serif-medium', small Arial and small Verdana were grouped as 'sans serif-small', medium Arial and medium Verdana were grouped as 'sans serif-medium', small Lexie Readable and small OpenDyslexic were grouped as 'dyslexia-optimised-small', and medium Lexie Readable and medium OpenDyslexic were grouped as 'dyslexia-optimised-medium'.

All data were entered into a MS Excel spreadsheet and then analysed using the SPSS statistical program. All data were visually analysed using histogram to check for normal distribution. Outliers for number of fixations, average fixation duration and reading time were adjusted using *winsorization* technique (DeCator, 2015). In this technique, extreme values below or above than *Mean* ± 2*SD* were adjusted.

3.3.3 Results

The majority of data for reading performance and eye gaze behaviour were normally distributed and heterogeneous, therefore they were analysed with parametric tests, specifically analysis of variance (ANOVA). Non-normality and heterogenous data appear to affect the result of ANOVA, however the test is robust enough when the size of samples are equal for all participants groups (Blanca, Alarcón, Arnau, Bono, & Bendayan, 2017; Field, 2013; Glass, Peckham, & Sanders, 1972; Laerd Statistics, 2015; Schmider, Ziegler, Danay, Beyer, & Bühner, 2010) which is the case in this data set.

A 3 x 3 x 2 mixed ANOVA was used, comprising 3 participant groups (non-dyslexic, mild dyslexic and moderate dyslexic), 3 typefaces (sans serif, serif and dyslexia-optimised) and 2 font sizes (small and medium). The dependent variables were number of fixations, average fixation duration and text reading time. Bonferroni-corrected post-hoc tests were conducted to investigate specific differences uncovered by the analyses. An alpha significance level of .05 was used in all statistical tests in this study. As well as the significant level, effect size is an objective and (usually) standardized measure of the magnitude of an observed effect (Field, 2013). It indicates how strongly two or more variables are related or how large the difference is between groups. According to Levine and Hullett (2002), partial eta squared is the most frequently reported measure of effect size for ANOVA. To estimate the magnitude of the effect size; partial eta squared of .01 is considered small, .06 is considered medium and .14 is considered large (Draper, 2018).

Since comprehension scores were used to confirm that participants had read the text, Chi-square tests were used to investigate any relationships between the comprehension scores and the three participant groups. For the preference measures, non-parametric tests were used because of the violation in the assumption of normality in the preference data. Therefore, median and interquartile range were used to describe the data.

3.3.3.1 Eye Gaze Behaviour

a. Number of Fixations

Table 3.5 shows the descriptive statistics for the number of fixations for the three participant groups. The number of fixations were normally distributed except for two groups: moderate dyslexic participants with dyslexia-optimised-medium font, $p = .02$; and moderate dyslexic participants with serif-medium font, $p = .01$, as assessed by Shapiro-Wilk's test of normality. There were no violations on homogeneity of variances, as assessed by Levene's test for equality of variances. Mauchly's test of sphericity indicated that the assumption of sphericity was met. The results of the ANOVA are summarised in Table 3.6.

Table 3.5 Descriptive statistics of number of fixations for all participant groups

	Non-Dyslexic		**Mild Dyslexic**		**Moderate Dyslexic**	
	M	***SD***	***M***	***SD***	***M***	***SD***
Sans serif-small	0.81	0.16	0.83	0.11	0.92	0.16
Sans serif-medium	0.87	0.17	0.90	0.14	1.12	0.23
Serif-small	0.76	0.14	0.80	0.15	0.90	0.20
Serif-medium	0.88	0.20	0.93	0.18	1.11	0.25
Dyslexia-optimised-small	0.86	0.17	0.84	0.17	0.95	0.22
Dyslexia-optimised-medium	0.86	0.18	0.95	0.17	1.10	0.24

Note. Measures count per word.

Table 3.6 Summary of three-way mixed ANOVA on number of fixations

Effects		***F***	***p***	η_p^2
Main	Group	$F(2, 33) = 3.82$	.03	.19
	Typeface	$F(2, 66) = 2.79$	ns	.08
	Font Size	$F(1, 33) = 62.29$	< .001	.65
Two-way interaction	Group × Typeface	$F(4, 66) = 0.39$	ns	.02
	Group × Font Size	$F(2, 33) = 5.97$	.01	.27
	Typeface × Font Size	$F(2, 66) = 4.15$	.02	.11
Three-way interaction	Group × Typeface × Font Size	$F(4, 66) = 1.18$	ns	.07

Note. 'ns' indicates not statistically significant.

There was a significant main effect for Group with large effect size, $F(2, 33) = 3.82$, $p = .03$, $\eta_p^2 = .19$. This is illustrated in Figure 3.8. A post-hoc analysis showed that moderate dyslexic participants ($M = 1.02$, $SD = 0.23$) had significantly higher number of fixations than non-dyslexic participants ($M = 0.84$, $SD = 0.17$), while the other differences were not significant.

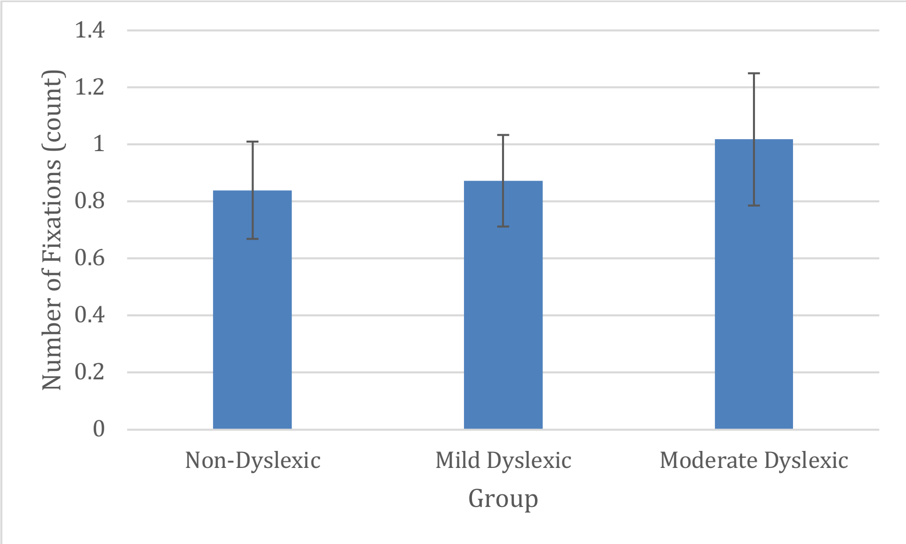


Figure 3.8 Post-hoc of Group on number of fixations

There was also a significant main effect for Font Size with large effect size, $F(1, 33) = 62.29$, $p < .001$, $\eta_p^2 = .65$ where medium font size ($M = 0.97$, $SD = 0.22$) had significantly higher number of fixations than small font size ($M = 0.85$, $SD = 0.17$).

There was a significant two-way interaction of Group and Font Size with large effect size, $F(2, 66) = 5.97$, $p = .01$, $\eta_p^2 = .65$. Figure 3.9 illustrates this effect. Post-hoc analysis indicated that all groups had a significantly higher number of fixations when read with medium font size compared to small font size ($p_{NonDyslexic} = .02$, $p_{MildDyslexic}$ and $p_{ModerateDyslexic} < .001$). Furthermore, moderate dyslexic participants had significant higher number of fixations than non-dyslexic participants when reading with medium font size ($p = .01$).

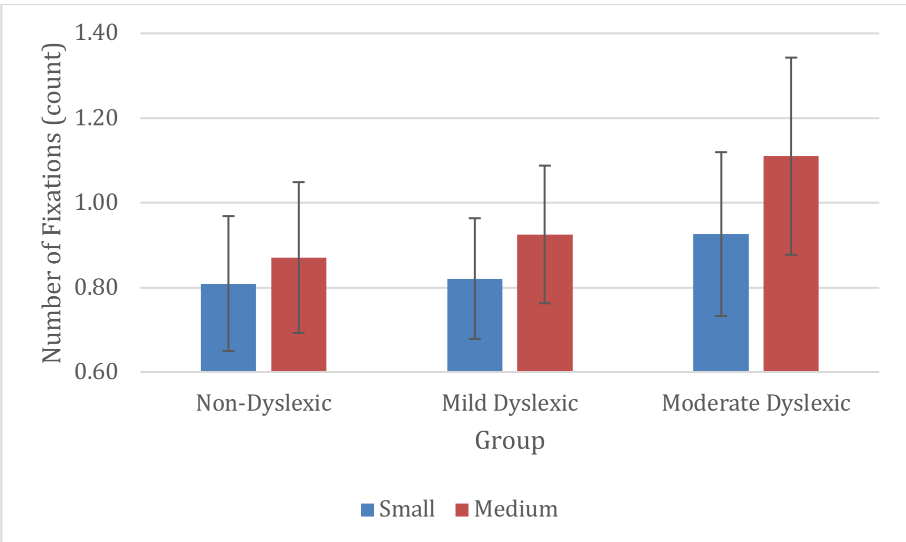


Figure 3.9 Interaction of Group and Font Size on number of fixations

There was a significant two-way interaction between Typeface and Font Size with moderate to large effect size, $F(2, 66) = 4.15$, $p = .02$, $\eta_p^2 = .11$. Figure 3.10 illustrates this effect. Post-hoc analysis indicated that all typefaces had significant higher number of fixations when read with medium than small font size (all three comparisons $p < .001$) and reading with dyslexia-optimised-small font had a significantly higher number of fixations compared to serif-small font ($p < .01$).

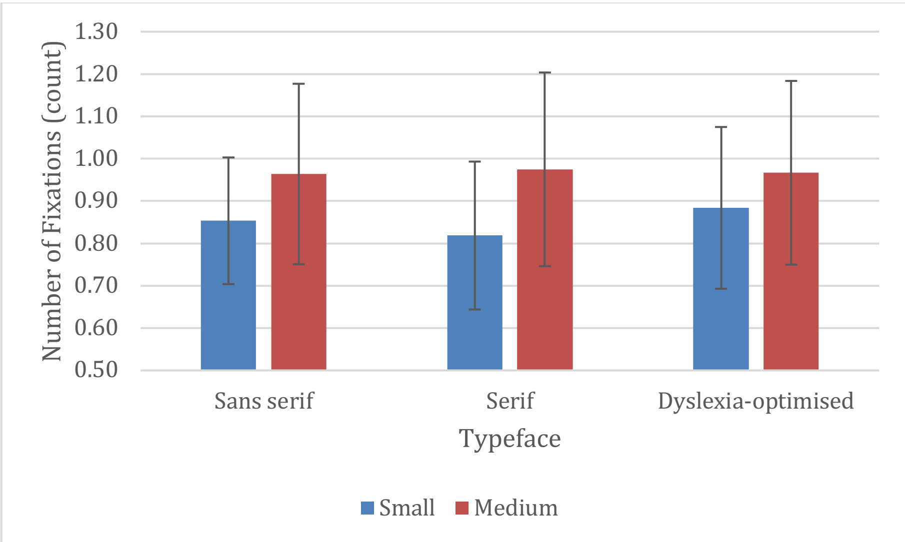


Figure 3.10 Interaction of Typeface and Font Size on number of fixations

b. Average Fixation Duration

Table 3.7 shows the descriptive statistics for the average fixation duration for the three participant groups. Average fixation duration were normally distributed except for one group: mild dyslexic participants with sans serif-small, $p = .02$, as assessed by Shapiro-Wilk's test of normality. There were no violations on homogeneity of variances, as assessed by Levene's test for equality of variances. Mauchly's test of sphericity indicated that the assumption of sphericity was violated, $X^2(2) = 7.78$, $p = .02$, therefore degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ($\varepsilon = .91$). The results of the ANOVA are summarised in Table 3.8.

Table 3.7 Descriptive statistics of average fixation duration for all participant groups

	Non-Dyslexic		Mild Dyslexic		Moderate Dyslexic	
	M	***SD***	***M***	***SD***	***M***	***SD***
Sans serif-small	3.65	0.63	3.57	0.37	4.08	0.58
Sans serif-medium	3.16	0.44	3.14	0.38	3.46	0.52
Serif-small	3.67	0.60	3.48	0.45	4.06	0.72
Serif-medium	3.29	0.49	3.14	0.33	3.40	0.46
Dyslexia-optimised-small	3.49	0.46	3.42	0.38	3.73	0.70
Dyslexia-optimised-medium	3.08	0.38	3.10	0.30	3.23	0.40

Note. Measures in milliseconds (ms) per word.

Table 3.8 Summary of three-way mixed ANOVA on average fixation duration

Effects		***F***	***p***	η_p^2
Main	Group	$F(2, 33) = 1.94$	ns	.11
	Typeface	$F(2, 66) = 21.57$	< .001	.40
	Font Size	$F(1, 33) = 150.03$	< .001	.82
Two-way interaction	Group × Typeface	$F(4, 66) = 3.47$	.01	.17
	Group × Font Size	$F(2, 33) = 3.26$	ns	.17
	Typeface × Font Size	$F(1.8, 66) = 1.86$	ns	.05
Three-way interaction	Group × Typeface × Font Size	$F(3.6, 66) = 0.74$	ns	.04

Note. 'ns' indicates not statistically significant.

There was a significant main effect for Typeface with large effect size, $F(2, 66) = 21.57$, $p < .001$, $\eta_p^2 = .40$. This is illustrated in Figure 3.11. A post-hoc analysis showed that dyslexia-optimised ($M = 3.34$, $SD = 0.50$) had significantly shorter fixation duration than the serif ($M = 3.51$, $SD = 0.59$) and sans serif ($M = 3.51$, $SD = 0.58$).

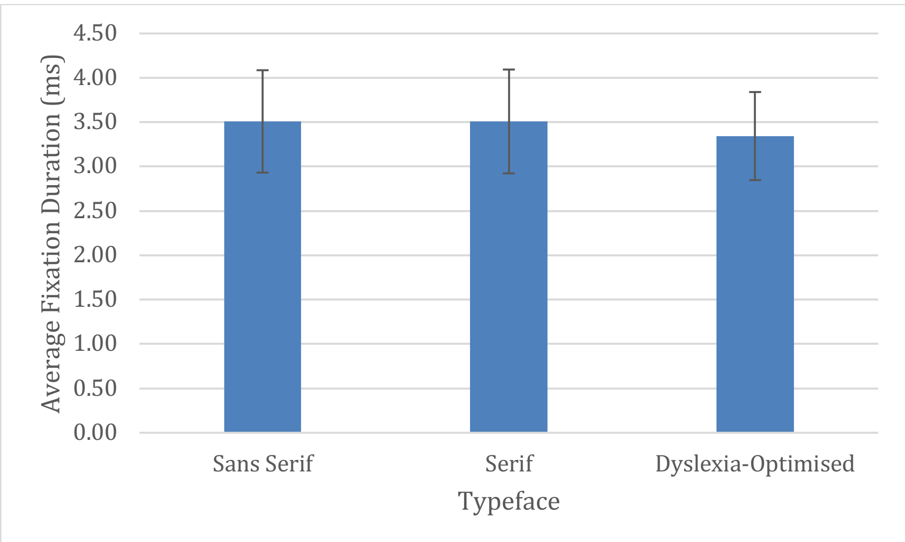


Figure 3.11 Interaction of Typeface on average fixation duration

There was also a significant main effect for Font Size with large effect size, $F(1, 33) = 150.03$, $p < .001$, $\eta_p^2 = .82$ where reading with medium font size ($M = 3.22$, $SD = 0.42$) had significantly shorter fixation duration than small font size ($M = 3.69$, $SD = 0.58$).

There was a significant two-way interaction for Group and Typeface with large effect size, $F(4, 66) = 3.47$, $p = .01$, $\eta_p^2 = .17$. Figure 3.12 illustrates this effect. Post-hoc analysis indicated that non-dyslexic participants had significantly shorter fixation duration with dyslexia-optimised in comparison with serif typefaces ($p < .01$). In addition, moderate dyslexic participants had significant shorter fixation duration with dyslexia-optimised compared to both serif and sans serif (both $p < .001$).

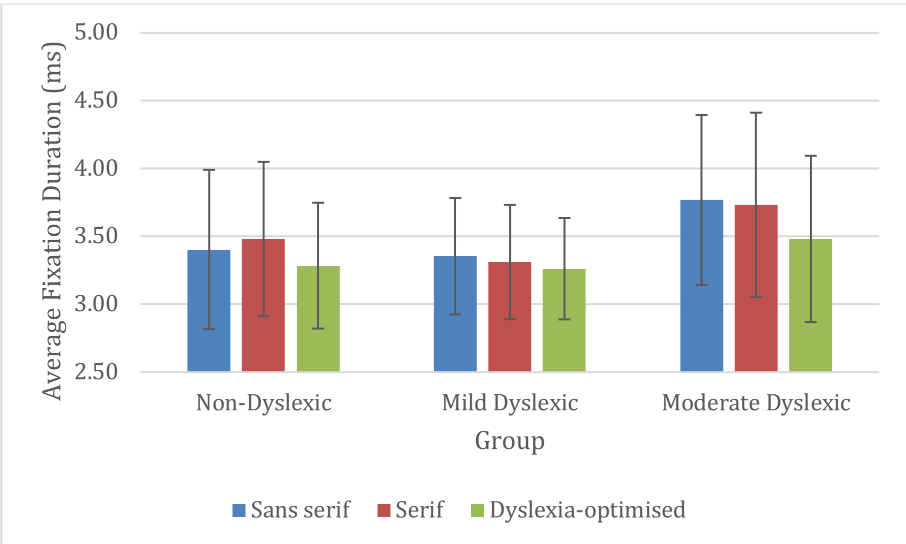


Figure 3.12 Interaction of Group and Typeface on average fixation duration

3.3.3.2 Reading Performance

a. Text Reading Time

Table 3.9 shows the descriptive statistics for the text reading time for the three participant groups. Text reading time were normally distributed, as assessed by Shapiro-Wilk's test of normality. The assumptions on homogeneity of variances was met except for Sans Serif-Medium ($F = 4.23$, $p = .02$), as assessed by Levene's test for equality of variances. Mauchly's test of sphericity indicated that the assumption of sphericity was not violated. The results of the ANOVA are summarised in Table 3.10.

Table 3.9 Descriptive statistics of text reading time for all participant groups

Font	Non-Dyslexic		Mild Dyslexic		Moderate Dyslexic	
	M	***SD***	***M***	***SD***	***M***	***SD***
Sans serif-small	233.32	61.72	229.28	42.37	288.94	68.78
Sans serif-medium	225.74	62.00	225.27	42.18	308.15	91.90
Serif-small	226.94	69.07	222.34	49.45	290.47	89.38
Serif-medium	238.30	75.42	237.50	49.64	304.17	87.09
Dyslexia-optimised-small	241.03	68.97	226.95	50.54	279.45	93.27
Dyslexia-optimised-medium	217.93	60.80	237.82	41.79	288.97	81.61

Note. Measures in milliseconds (ms) per word.

Table 3.10 Summary of three-way mixed ANOVA on text reading time

Effects		*F*	*p*	η_p^2
Main	Group	$F(2, 33) = 3.88$	.03	.19
	Typeface	$F(2, 66) = 0.72$	ns	.02
	Font Size	$F(1, 33) = 1.62$	ns	.05
Two-way interaction	Group × Typeface	$F(4, 66) = 1.23$	ns	.07
	Group × Font Size	$F(2, 33) = 32.36$	ns	.13
	Typeface × Font Size	$F(2, 66) = 1.68$	ns	.05
Three-way interaction	Group × Typeface × Font Size	$F(4, 66) = 1.28$	ns	.07

Note. 'ns' indicates not statistically significant.

There was a significant main effect for Group with large effect size, $F(2, 33) = 3.88$, $p = .03$, $\eta_p^2 = .19$. Figure 3.13 illustrates this effect. However, post-hoc analysis indicated there were no significant differences on reading time between non-dyslexic participants ($M = 230.54$, $SD = 64.64$), mild dyslexic participants ($M = 229.86$, $SD = 44.90$) and moderate dyslexic participants ($M = 293.36$, $SD = 83.25$). Based on Figure 3.10, moderate dyslexic participants showed higher mean reading time and a larger standard deviation in their reading times than non-dyslexic and mild dyslexic participants.

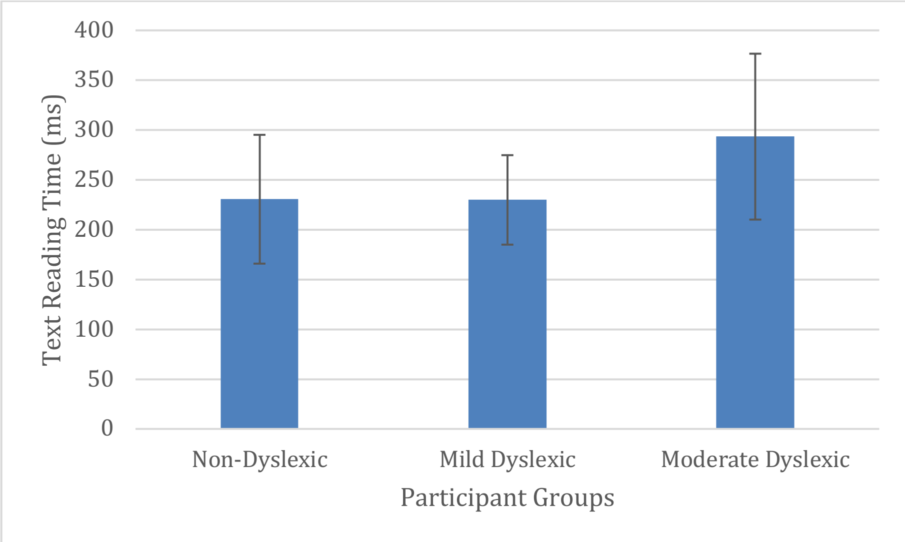


Figure 3.13 Mean of Text Reading Time for all groups

b. Comprehension Scores

Participants answered a total of 36 comprehension questions across the 12 texts. The mean comprehension score for moderate dyslexic participants (*M* = 25.75, *SD* = 3.91) was slightly lower than mild dyslexic (*M* = 27.17, *SD* = 3.76) and non-dyslexic (*M* = 26.17, *SD* = 2.33). To investigate the comprehension scores, a Chi-square test of independence was conducted between comprehension scores and the three participants groups, three categories of typefaces and two font sizes. There was no relationship between comprehension scores with Group ($X^2(2) = 0.23$, $p = .89$), Typeface ($X^2(2) = 1.02$, $p = .60$), or Font Size ($X^2(1) = 2.72$, $p = .10$).

3.3.3.3 Participants' Preferences and Opinions

a. Ease of Reading of Different Typefaces and Different Sized Fonts

Figure 3.14 shows the median and interquartile range ratings given for ease of reading for all combinations of typeface and font size. Note that lower ratings of ease of reading indicate greater difficulty of reading while higher ratings indicate greater ease of reading. It can be seen that Arial-medium size had the highest rating (*M* = 5.67, *SD* = 1.07, *Mdn* = 6.00) of ease of reading so was rated the easiest to read, while OpenDyslexic-small had the lowest rating (*M* = 3.41, *SD* = 1.82, *Mdn* = 2.83) of ease of reading, so was rated the most difficult. Complete descriptive statistics of rating given for ease of reading for all combinations of typefaces and font size can be found in Appendix I.

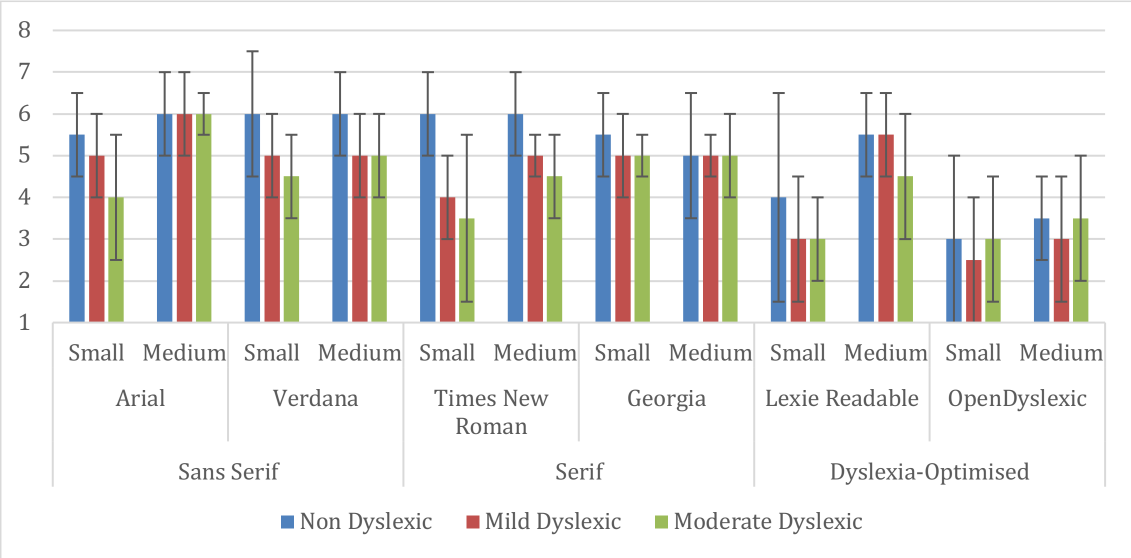


Figure 3.14 Ease of reading ratings for different typefaces and font sizes

To investigate whether there were any differences in ratings of ease of reading for each typeface and font size combination for each group of participants, Friedman tests were used. Pairwise comparisons were performed using Wilcoxon Signed Rank tests to investigate whether there were differences in the rating of ease of reading between different typefaces and different font sizes within each participant group.

For the non-dyslexic group, ratings of ease of reading were statistically different between the three typefaces, $X^2(2, n = 12) = 11.46, p < .001$. Figure 3.15 shows the median ease of reading ratings and IQR for the three typefaces. Post-hoc analysis revealed that participants rated the Sans Serif typefaces (*Mdn* = 5.50, *IQR* = 1.38) significantly easier to read than the Dyslexia-Optimised typefaces (*Mdn* = 3.75, *IQR* = 2.88), $z = 2.57$, $p = .01$, $r = .7$, and the Serif typefaces (*Mdn* = 5.50, *IQR* = 1.38) significantly easier to read than the Dyslexia-Optimised typefaces (*Mdn* = 3.75, *IQR* = 2.88), $z = -2.64$, $p = .01$, $r = -.76$. However, there was no statistically significant difference in the ratings of ease of reading between Sans Serif and Serif typefaces, $z = .93$, $p = .35$, $r = .29$. For font size, a Wilcoxon Signed Rank test determined that there was no statistically significant difference in the ratings of ease of reading between the small font size (*Mdn* = 5.25, *IQR* = 1.38) and the medium font size (*Mdn* = 5.25, *IQR* = 1.75), $z = -.41$, $p = .69$, $r = -.12$.

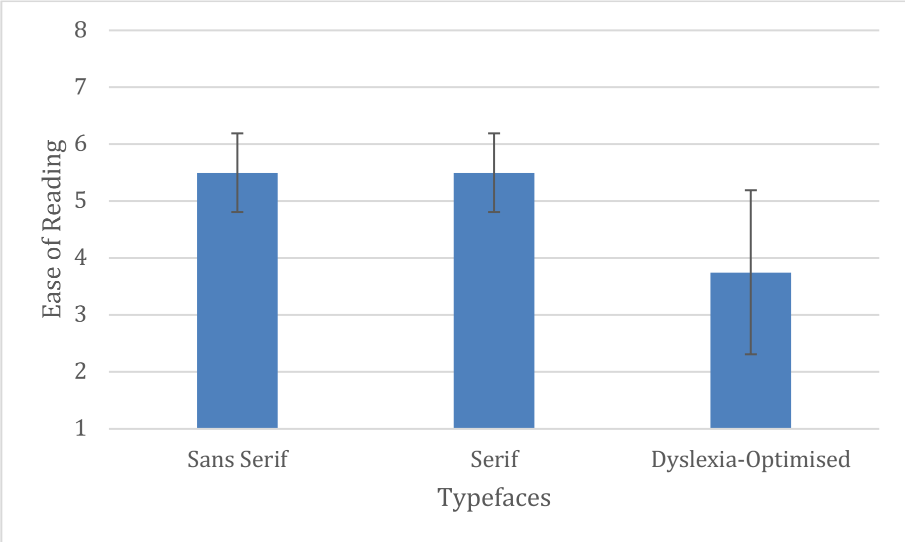


Figure 3.15 Ease of reading ratings of non-dyslexic participants for sans serif, serif and dyslexia-optimised typefaces

For the mild dyslexic group, ratings of ease of reading were statistically different between the three typefaces, $X^2(2, n = 12) = 8.58$, $p = .01$. Figure 3.16 shows the median ease of reading ratings and IQR for the three typefaces. Post-hoc analysis revealed that participants rated the sans serif typefaces (*Mdn* = 5.25, *IQR* = 1.38) significantly easier to read than the dyslexia-optimised typefaces (*Mdn* = 3.25, *IQR* = 3.13), $z = 2.02$, $p = .04$, $r = .58$ and the serif typefaces (*Mdn* = 4.75, *IQR* = 0.88) significantly easier than dyslexia-optimised typefaces (*Mdn* = 3.25, *IQR* = 3.13), $z = -2.21$, $p = .03$, $r = -.64$. However, there was no statistically significant difference in the ratings of ease of reading between sans serif and serif typefaces, $z = -.66$, $p = .51$, $r = -.19$. For font size, a Wilcoxon Signed Rank test determined that there was no statistically significant difference in the ratings of ease of reading between the small font size (*Mdn* = 4.50, *IQR* = 1.25) and medium font size (*Mdn* = 5.00, *IQR* = 1.25), $z = -1.38$, $p = .17$, $r = -.40$.

For the moderate dyslexic group, differences in ratings of ease of reading between the three typefaces were not statistically significant, $X^2(2, n = 12) = 5.32$, $p = .07$. For font size, a Wilcoxon Signed Rank test determined that medium font (*Mdn* = 5.00, *IQR* = 1.38) was significantly rated easier to read than small font (*Mdn* = 3.50, *IQR* = 1.88) to, $z = -2.85$, $p < .001$, $r = -.82$.

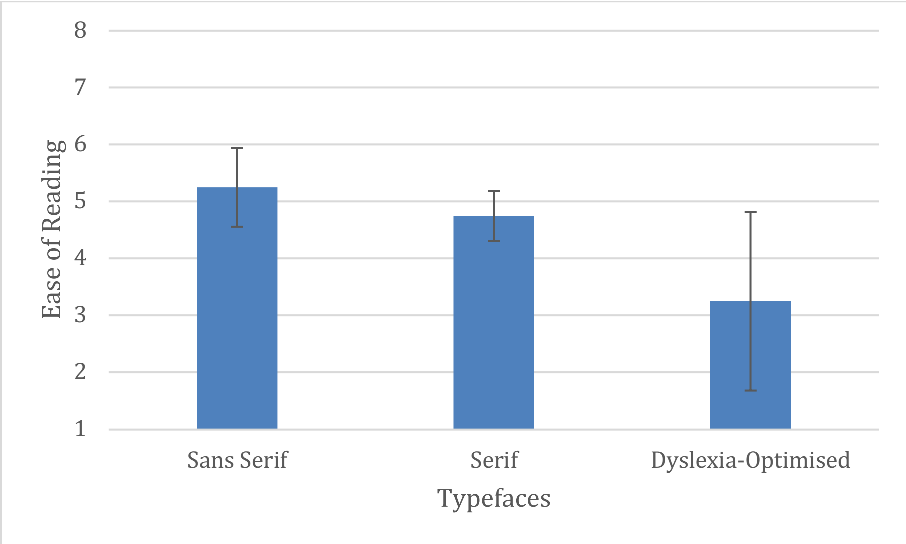


Figure 3.16 Ease of reading ratings of mild dyslexic participants for sans serif, serif and dyslexia-optimised typefaces

b. Preferred Typeface and Font Size Combinations

Table 3.11 shows the frequency of preferred typeface for each participant group. To investigate participants' preferred combination of typeface and font size, firstly a Chi-square test of independence was conducted between the three participant groups and typeface on their stated preferred combination. There was no statistically significant association between participant groups and the typeface of their preferred combination, $X^2(4, n = 36) = 2.81, p = .59$.

Table 3.11 Distribution of preferences for typeface by the three participant groups

Participant Group	N	Typeface		
		Sans Serif	**Serif**	**Dyslexia-Optimised**
Non-Dyslexic	12	8	3	1
Mild Dyslexic	12	6	4	2
Moderate Dyslexic	12	4	6	2
TOTAL	**36**	**18**	**13**	**5**

However, when a Chi-square likelihood test was conducted on preference for typeface for all participants together, there was a significant difference in preference for typeface, $X^2(2, \text{n} = 36) = 7.17, p = .03$. From Table 3.11 it can be seen that 50% the participants (18/36) preferred the sans serif typefaces, 36% (13/36) preferred the serif typefaces and only 14% (5/36) preferred the dyslexia-optimised typefaces.

Table 3.12 shows the frequency of preferred font size for each participant group. By continuing the investigation described above, a second Chi-square test of independence was conducted between the three participant groups and font size on their stated preferred combination. There was no statistically significant association between participant groups and the font size of their preferred combination, $X^2(2, n = 36) = 0.22, p = .89$.

Table 3.12 Distribution of preferences for font size by the three participant groups

Participant Group	**N**	**Font Size**	
		Small	**Medium**
Non-Dyslexic	12	6	6
Mild Dyslexic	12	7	5
Moderate Dyslexic	12	6	6
TOTAL	**36**	**19**	**17**

In addition, when a Chi-square likelihood test was conducted on preference for font size for all participants together, there was no significant difference in preference for font size, $X^2(1, n = 36) = 2.06, p = .15$. From Table 3.12 it can be seen that almost 53% of the participants (19/36) preferred the small font and around 47% (17/36) preferred the medium font.

c. Previous Engagement on Dyslexia-Optimised Typefaces

Table 3.13 shows the distribution of participants and whether they have encountered either of the dyslexia-optimised typefaces before the study. To investigate participants' previous engagement with the typefaces, a Chi-square test of independence was conducted between the number of participants in each group and their previous encounters with each dyslexia-optimised typeface. There was no statistically significant association between the groups and the number of participants who had encountered Lexie Readable typeface, $X^2(2, n = 36) = 1.83, p = .40$. Neither was there a statistically significant association between the groups and the number of participants who have encountered OpenDyslexic typeface, $X^2(2, n = 36) = .89, p = .64$.

Table 3.13 Numbers of participants who had previously encountered on dyslexia optimised typefaces

	Lexie Readable		**OpenDyslexic**	
Participant Group	Yes	No	Yes	No
Non-Dyslexic	4	8	2	10
Mild Dyslexic	5	7	4	8
Moderate Dyslexic	2	10	3	9
TOTAL	**11**	**25**	**9**	**27**

However, when a Chi-square likelihood test was conducted on their previous encounters with dyslexia-optimised typefaces for all participants together, there was a significant difference in number of participants who had encountered Lexie Readable typeface, $X^2(1, n = 36) = 5.44, p = .02, r = .39$. A similar result was found for OpenDyslexic, $X^2(1, n = 36) = 9.00, p < .001, r = .50$. From Table 3.13 it can be seen that around 69% of the participants (25/36) have never encountered Lexie Readable and 75% (27/36) of participants have never encountered OpenDyslexic before participating in the study. It can be seen that moderate dyslexic participants had the lowest previous engagement with Lexie Readable and slightly above than non-dyslexic participants with OpenDyslexic. On average, around 72% of participants have never encountered dyslexia-optimised typefaces.

d. Dyslexic Participants Opinions on Dyslexia-Optimised Typefaces

To investigate the dyslexic participants' opinions of the two dyslexia-optimised typefaces, they were asked to rate ease of reading of the typefaces in comparison with other standard typefaces and also to rate their interest to use the typefaces on 7-Likert items. Table 3.14 shows the descriptive statistics of the ratings. The total number of participants with dyslexia was 24, however one dyslexic participant did not complete these ratings. Lower rating for 'Easier to read than standard typefaces' indicates the typeface are difficult to read compared to standard typefaces, while lower rating for 'Interested to use' indicates participants were not interested to use the typefaces in their computer. From Table 3.14, it can be seen participants with dyslexia rated OpenDyslexic as difficult to read compared to standard typefaces and they had less interest in using the typeface in their computer. However, Lexie Readable was given a better rating for both questions.

Table 3.14 Descriptive statistics of dyslexic participants' opinions on dyslexia-optimised typefaces

	***M* (*SD*)**	**95% CI**	***Mdn***	***IQR***
Easier to read than standard typefaces				
Lexie Readable	3.48 (1.83)	[2.69, 4.27]	3.00	3.00
OpenDyslexic	2.70 (1.96)	[1.85, 3.54]	2.00	3.00
Interested to use				
Lexie Readable	3.26 (1.79)	[2.49, 3.03]	3.00	3.00
OpenDyslexic	2.48 (1.93)	[1.64, 3.31]	2.00	2.00

Figure 3.17 shows distribution of dyslexic participants' ratings on whether the typefaces were easier to read compared with other standard typefaces. To investigate whether participants rated the typefaces as easier to read, Chi-square likelihood tests were conducted on Lexie Readable and OpenDyslexic. For Lexie Readable, the distribution of ratings was not different from a random distribution, $X^2(6, n = 23) = 7.13$, $p = .31$. For OpenDyslexic, the distribution of ratings was significantly different from a random distribution, $X^2(6, n = 23) = 14.44$, $p = .03$. From Figure 3.17, 56.5% of the participants (13/23) rated OpenDyslexic as 1 or 2 on the 7-point scale, so definitely no easier to read or only one level above that, whereas only 13.0% of participants (3/23) rated it 6 or 7, so definitely easier to read or one level below that.

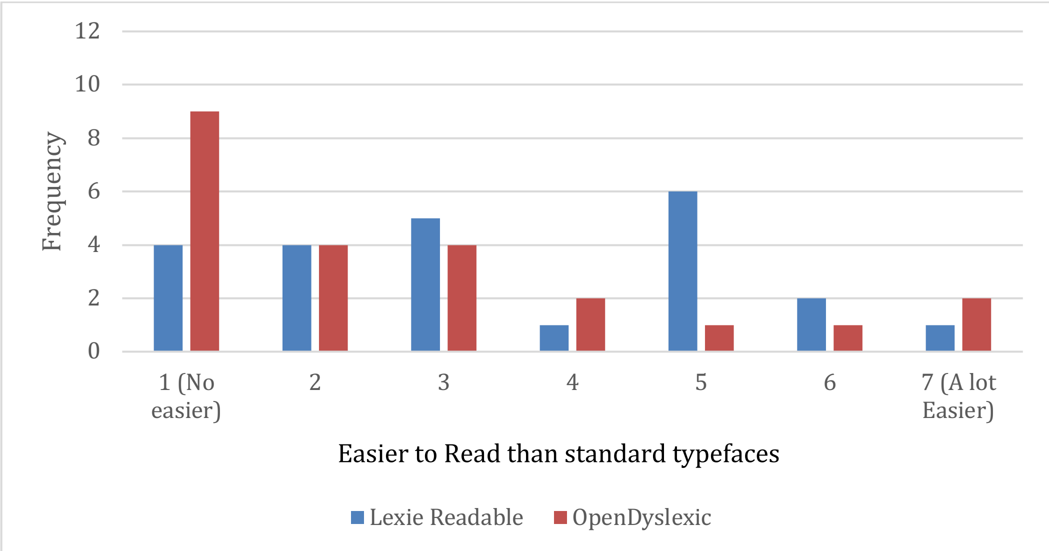


Figure 3.17 Distribution of ratings whether the dyslexia-optimised typefaces were easier to read than standard typefaces

To investigate whether there were any differences between dyslexic groups on their opinions of whether the dyslexia-optimised typefaces were easier to read compared to standard typefaces, Kruskal-Wallis H tests were conducted comparing the ratings for mild and moderate dyslexic participants. No statistically significant differences were found for either typefaces.

To investigate whether the median of the ratings of both typefaces are easier to read compared with other standard typefaces differs from the midpoint of the 7-Likert items (i.e. a rating of 4), a Wilcoxon Signed Rank tests were conducted. For Lexie Readable, the median rating for easier to use ($Mdn = 3$, $IQR = 3$) was not statistical significantly different from the midpoint, $z = -1.37$, $p = .17$, $r=-.29$. For OpenDyslexic, the median of rating for easier to use (*Mdn* = 2, *IQR* = 3) was statistically significant lower from the midpoint, *z* = -2.57, *p* = .01, *r* = -.54.

Figure 3.18 shows distribution of dyslexic participants' ratings of interest to use the dyslexia-optimised typefaces on their computer. To investigate the distribution of participants' ratings of their interest to use the typefaces, Chi-square likelihood test with Lexie Readable and OpenDyslexic were conducted. For Lexie Readable, the test showed that the distribution of dyslexic participants that rated on their interest to use the typeface was not different from a random distribution, $X^2(6, n = 23) = 4.70$, $p = .58$. For OpenDyslexic, the test showed that the distribution of dyslexic participants that rated their interest to use the typeface was different from a random distribution, $X^2(5, n = 23) = 14.30$, $p = .01$. From Figure 3.18, 65.2% of the participants (15/23) rated OpenDyslexic as 1 or 2 on the 7-point scale, so definitely not interested to use or only one level above that, whereas 13% of the participants (3/23) rated it 6 or 7, so definitely interested to use or one level below that.

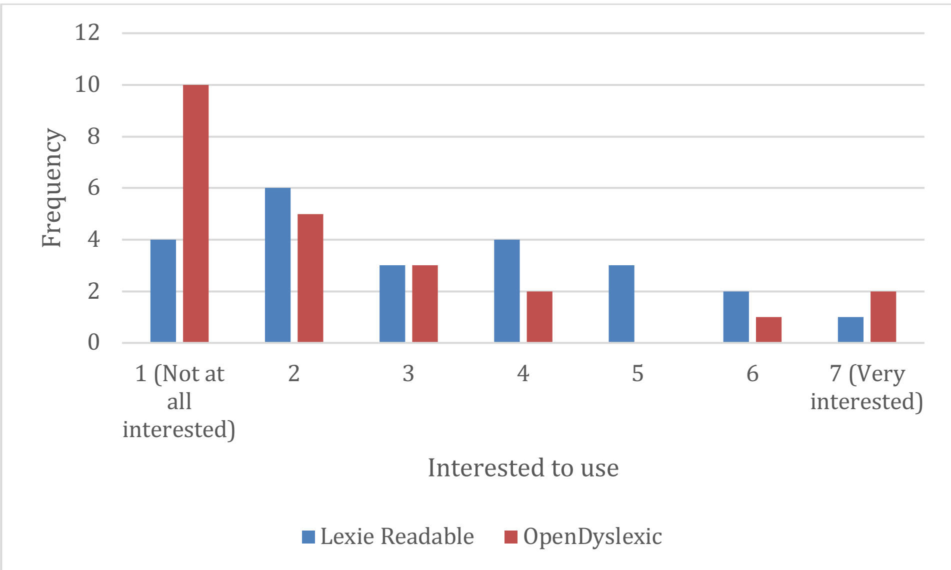


Figure 3.18 Distribution of ratings of interest in using dyslexia-optimised typefaces

To investigate whether there were any differences between dyslexic groups on their interest to use the dyslexia-optimised typefaces, Kruskal-Wallis H tests were conducted. No statistically significant differences were found for either typefaces.

To investigate whether the median of the ratings of interest to use both typefaces differs from the midpoint of the 7-Likert items (i.e. a rating of 4), a Wilcoxon Signed Rank tests were conducted. For Lexie Readable, the median rating for interest to use (*Mdn* = 3, *IQR* = 3) was not statistically significant different from the midpoint, *z* = -1.90, *p* = .06, *r* = -

.40. For OpenDyslexic, the median of rating for interest to use (*Mdn* = 2, *IQR* = 2) was statistical significantly lower from the midpoint, $z = -2.78$, $p = .01$, $r = -.58$. Thus, dyslexic participants are not interested in using OpenDyslexic.

e. Comfort with Dyslexia-Optimised Typefaces

Participants' comfort in using the two dyslexia-optimised typefaces was investigated differently, depending on whether the participants were dyslexic or not. Non-dyslexic participants were asked about their comfort in receiving material presented in the typefaces to read, while dyslexic participants were asked about their comfort in submitting materials to teachers or colleagues using the typefaces. In both instances, participants were asked to rate their level of comfort on a scale from 1 (not comfortable) to 7 (very comfortable).

Table 3.15 shows the descriptive statistics for ratings on their comfort with the typefaces. Lower ratings indicate the participants were not comfortable to receive or submit materials using the typeface. It can be seen that dyslexic participants gave lower rating on submitting materials to their teachers or colleagues with the typefaces than the non-dyslexic participants gave in receiving such materials.

Table 3.15 Descriptive statistics for dyslexics' comfortability on dyslexia-optimised typefaces

	***M* (*SD*)**	**95% CI**	***Mdn***	***IQR***
Non-Dyslexic				
Lexie Readable	4.58 (2.31)	[3.11, 6.05]	5.00	5.00
OpenDyslexic	3.42 (1.93)	[2.19, 4.64]	3.00	4.00
Mild Dyslexic				
Lexie Readable	3.33 (2.10)	[2.00, 4.67]	3.00	4.00
OpenDyslexic	2.33 (1.72)	[1.24, 3.43]	1.50	3.00
Moderate Dyslexic				
Lexie Readable	3.25 (2.18)	[1.87, 4.63]	3.00	4.00
OpenDyslexic	3.00(2.52)	[1.40, 4.60]	1.50	5.00

In order to analyse participants' comfort on dyslexia-optimised typefaces, separate analyses were conducted on dyslexic and non-dyslexic groups because they were asked two different questions. Figure 3.19 shows distribution of ratings of non-dyslexic participants comfort in receiving materials in each of the dyslexia-optimised typefaces. It can be seen that four non-dyslexic participants felt very comfortable (giving a rating of 7) to receive materials with Lexie Readable typeface, three of them felt very comfortable (giving rating of 6) to receive materials with OpenDyslexic. Figure 3.20 and Figure 3.21 shows distribution of ratings of comfort by mild and moderate dyslexic participants in submitting materials with the two dyslexia-optimised typefaces. It can be seen that seven dyslexic participants (3 mild dyslexic and 4 moderate dyslexic) felt not at all comfortable (giving rating of 1) to submit material in Lexie Readable typeface, twelve dyslexic participants (6 mild dyslexic and 6 moderate dyslexic) felt not at all comfortable (giving rating of 1) to submit materials in OpenDyslexic.

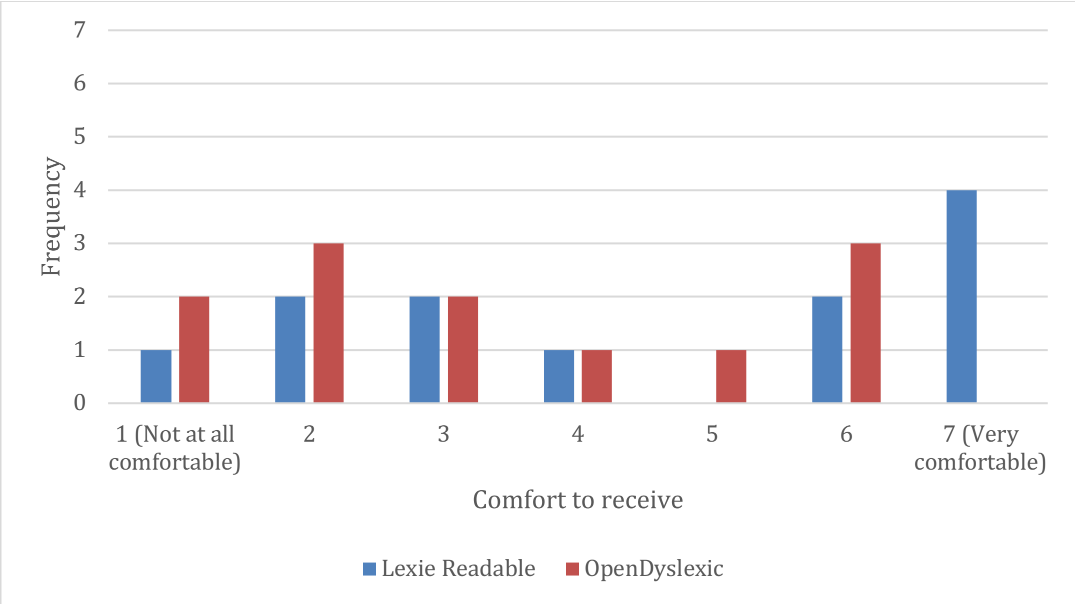


Figure 3.19 Distribution of comfort ratings by non-dyslexic participants on dyslexia-optimised typefaces

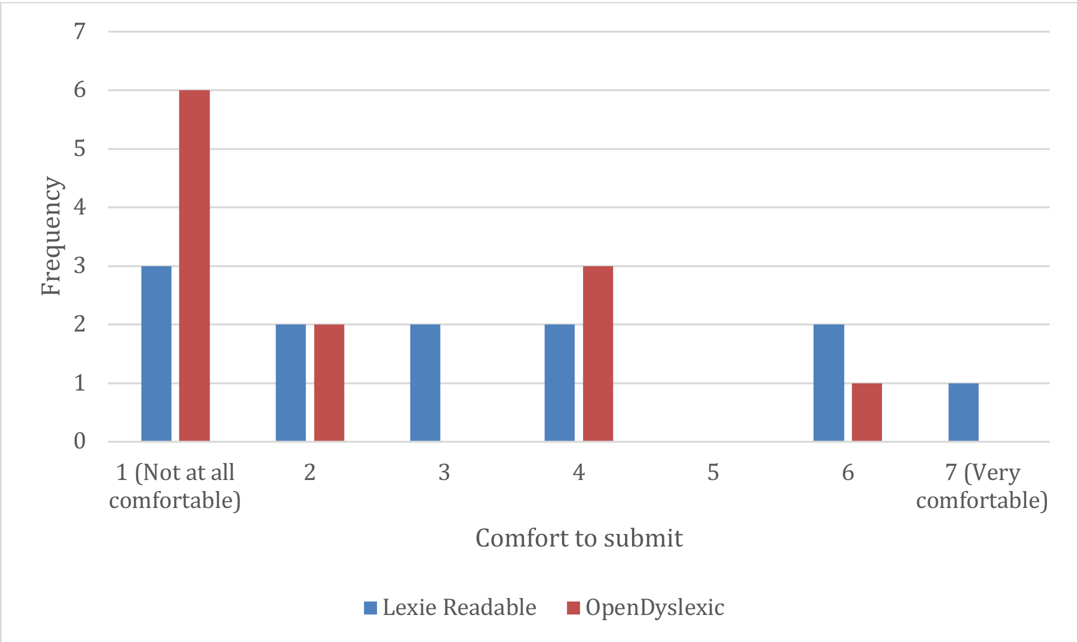


Figure 3.20 Distribution of comfort ratings of mild-dyslexic participants on dyslexia-optimised typefaces

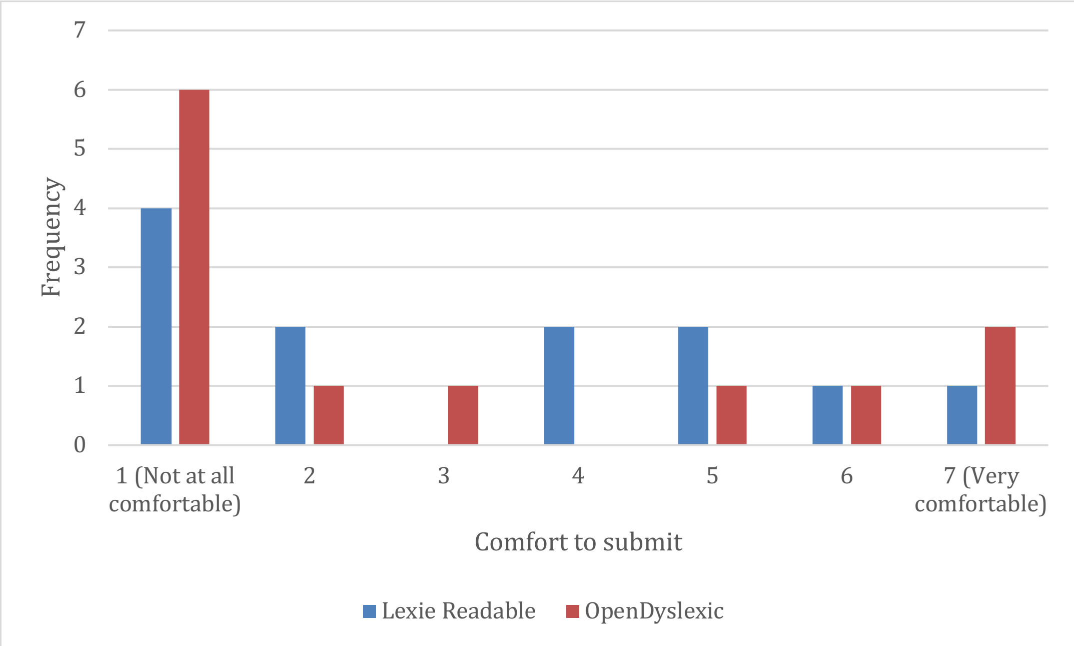


Figure 3.21 Distribution of comfort ratings of moderate dyslexic participants on dyslexia-optimised typefaces

To investigate the distribution of ratings of comfort by non-dyslexic participants, Chi-square likelihood tests were conducted on Lexie Readable and OpenDyslexic. Both tests showed that the distribution of ratings for participants' comfort in receiving materials using Lexie Readable and OpenDyslexic were not different from random distributions.

To investigate the distribution of comfort ratings for dyslexic participants (considered as one group) with the dyslexia-optimised typefaces, Chi-square likelihood tests were conducted on Lexie Readable and OpenDyslexic. The test showed that the distribution of comfort ratings for Lexie Readable was not different from a random distribution. However, there was a significant difference in the distribution of comfort ratings for OpenDyslexic from a random distribution, $X^2(6, n = 24) = 26.17, p = .00$. From Figure 3.20 and Figure 3.21, it can be seen that 50% of the dyslexic group (12/24) gave a rating of only 1 on the 7-Likert items, which means they were not at all comfortable to submit materials to colleagues or teachers using the typeface.

To investigate whether comfort of submitting materials with the typefaces differed between the two dyslexic groups (mild dyslexic and moderate dyslexic), Mann-Whitney U tests were conducted on each typeface. No statistically significant differences were found between the two dyslexic groups on their comfort with either typefaces.

In addition, Wilcoxon Signed Rank tests were conducted to investigate whether the median of the comfort ratings by dyslexic and non-dyslexic participants to submit materials using dyslexia-optimised typefaces differed from the midpoint of 4. For Lexie Readable ($Mdn_{NonDyslexic}$ = 5.00, $Mdn_{Dyslexic}$ = 3.00), there was no statistically significant difference from the midpoint. However, for OpenDyslexic ($Mdn_{NonDyslexic}$ = 3.00, $Mdn_{Dyslexic}$ = 2.00), the median rating for dyslexic participants was significantly lower than the midpoint, z = -2.71, p = .01, r = -.55. Thus, dyslexic participants are not comfortable in using OpenDyslexic.

3.3.4 Discussion and Conclusion

This study investigated the effects of typeface and font size on reading behaviour, reading performance and preference measures on detailed text reading from a computer screen. This study compared adult English native speakers with mild and moderate dyslexia and those without dyslexia. Participants were asked to complete an online dyslexia assessment checklist (the ARQ) and were categorised based on their severity of dyslexia into non-dyslexic, mild and moderate dyslexic. Overall, this study found that typeface had a significant effect on fixation duration, ease of reading, preferences and opinions, no significant effect on number of fixations and reading time. Font size had a significant effect on number of fixations, fixation duration and ease of reading, while no significant effect on reading time. Participant group had significant effect on number of fixations, no significant effect on fixation duration and reading time. Details for each of these aspects are discussed below.

3.3.4.1 Effects of Typeface

According to numerous guidelines for presenting text to people with dyslexia, sans serif is recommended as the typeface (British Dyslexia Association, 2018b; de Santana et al., 2012; Evett & Brown, 2005; Friedman & Bryen, 2007; Rello, 2014a; Rello & Baeza-Yates, 2015; Zarach, 2002), while some recommend to use mono-spaced (British Dyslexia Association, 2018b; de Santana et al., 2012; Rello, 2014a). It is important to note that most of the guidelines were created without any empirical research and only some of the guidelines incorporated dyslexia-optimised typefaces as part of their background research (Rello, 2014a; Rello & Baeza-Yates, 2015). Furthermore, a variety of studies on the performance of dyslexia-optimised typefaces for reading had been discussed in Chapter 2 (de Leeuw, 2010; Harley et al., 2013; Kuster et al., 2017; Marinus et al., 2016; Pijpker, 2013; Rello & Baeza-Yates, 2013, 2016; Wery & Diliberto, 2017).

Although sans serif was recommended for people with dyslexia, this present study found sans serif, serif and dyslexia-optimised typefaces had no significant effect on dyslexics' reading time as found in other studies (de Leeuw, 2010; Harley et al., 2013; Kuster et al., 2017; Pijpker, 2013; Rello & Baeza-Yates, 2013, 2016; Wery & Diliberto, 2017). However in a study with poor readers, Marinus et al. (2016) found out that dyslexia-optimised typeface (Dyslexie) was 7% more faster than reading with sans serif (Arial).

Similar to results observed by Rello & Baeza-Yates (2016), this study found that typeface had no significant effect on the number of fixations. Not only that, there was a significant interaction between Typeface and Font size, with participants having a significantly lower number of fixations when reading with Serif-Small font. This result may related to the sample being mainly university students. It is widely accepted to use Times or Times New Roman in 12pt (smaller than font sizes used in this present study) as part of many materials at the university. Both Times and Times New Roman are serif typefaces. Therefore, participants are quite familiar with Serif-Small font. In addition, Kaspar et al. (2015) found university students felt reading scientific abstracts with serif typefaces make the abstract more appealing, more interesting and more comprehensible.

This present study also found that dyslexia-optimised typefaces had significantly shorter fixation durations than sans serif and serif typefaces for all participants. However, while there was no main effect of typeface on the number of fixations, there was a significant interaction of typeface and font size, with the difference in number of fixations between small and medium font size being greatest for serif typefaces. In addition, when participants read dyslexia-optimised typeface in small size, they had higher number of fixations than reading with serif typefaces in small size. Furthermore, the typeface and font size had no significant difference on reading times. Overall, as dyslexia-optimised typefaces only produced better results in one (fixation duration) out of three objective measures (fixation number, fixation duration and reading time), dyslexia-optimised typefaces do not have better readability than other standard typefaces. Readability of a typeface improves when people make fewer fixations and shorter fixation duration (Rello & Baeza-Yates, 2015; Schneps, Thomson, Sonnert, et al., 2013)

This present study found that non-dyslexic and moderate dyslexic participants had significantly shorter fixation durations with dyslexia-optimised typefaces, a result which contradicts results found by Rello and Baeza-Yates (2016). In the Rello and Baeza-Yates (2016) study, non-dyslexic and dyslexic participants had longer fixation duration with a dyslexia-optimised typeface in comparison to sans serif typefaces. However, that study used only one dyslexia-optimised typeface, OpenDyslexic. Perhaps with a wider selection of dyslexia-optimised typefaces such as Lexie Readable and Dyslexie, different typeface effects would have been found.

Despite having shorter fixation durations, dyslexia-optimised typefaces were rated more difficult and least preferred compared to serif and sans serif typefaces. This contradictory result may have partly occurred because most participants were not familiar with dyslexia-optimised typefaces, so leading to beneficial effects towards familiar typefaces such as sans serif and serif (Grigorovich-Barsky & Belson, 2013; Marinus et al., 2016). In addition, the present study found an average of 72% of the participants had never encountered the dyslexia-optimised typefaces before participating in the study.

Dyslexic participants were asked to rate both dyslexia-optimised typefaces (OpenDyslexic and Lexie Readable) separately, they gave the OpenDyslexic typeface a rating significantly lower than median whereas Lexie Readable were rated neutral in terms of its readability versus standard typefaces, their interest and their comfort using the typeface. According to Hillier (2006), dyslexic readers dislike OpenDyslexic because they prefer light and uniform strokes. In addition, other studies have shown that OpenDyslexic did not lead to better reading performance (Harley et al., 2013; Rello & Baeza-Yates, 2013, 2016; Wery & Diliberto, 2017).

Dyslexic participants felt less comfortable to submit materials using dyslexia-optimised typefaces compared to non-dyslexic participants who were neutral and willing to accept materials in either dyslexia-optimised typefaces. Dyslexic participants perhaps have concerns about the perceptions made by non-dyslexic people when receiving materials in dyslexia-optimised typefaces. As found by De Leeuw (2010), people with dyslexia would rather not use dyslexic-optimised typefaces and are willing to cope with their difficulties by adapting to non-dyslexic.

3.3.4.2 Effects of Font Size

Previous research has found that reading using small font size is one of the problems for people with dyslexia (McCarthy & Swierenga, 2010). It is also important to identify the minimum size that is appropriate for dyslexic readers, as reading using a size lower than the minimum will significantly decrease reading speed and increase the number of fixations and their duration (O'Brien et al., 2005; Tinker & Paterson, 1955). The majority of the dyslexia guidelines recommend the use of font size between 12pt and 14pt (British Dyslexia Association, 2018b; de Santana et al., 2012; Evett & Brown, 2005;

Friedman & Bryen, 2007; Zarach, 2002). Based on empirical studies with eye tracking, a range between 18pt and 26pt was recommended for reading for both people with and without dyslexia (Rello, 2014a; Rello & Baeza-Yates, 2015; Rello et al., 2012b).

Other eye tracking studies on font size reported that participants had longer fixation duration with small font size compared to medium font size (Rello & Baeza-Yates, 2015; Rello, Pielot, et al., 2013). A similar result was found in this present study. However, the increase in fixation duration for the small font size was compensated by a significant reduction on the number of fixations, and the size had no effect on reading time. In addition, all groups of participants had a significant higher number of fixations with medium font size compared to small font size. According to Tinker & Paterson (1955), reading with smaller or bigger than optimal font size will increase the number of fixations. Therefore, medium font size used in this study probably was bigger than participants' optimal size for reading. Moreover, as the sample of participants were students in the university, they are familiar with 12pt size which is smaller than the font size used in this study.

Based on the participants' ratings on the preference measures, this study found that non-dyslexic and mild dyslexic participants had no difference in ratings of ease of reading for both font sizes. However, moderate dyslexic participants rated reading with medium font size as easier. In terms of their preferred combination to read, there was no significant difference between participants who chose small and medium font size. However Rello, Kanvinde, et al. (2012b), found that reading with Arial 14pt, 18pt, 22pt and 26pt showed around 64% of dyslexic participants preferred to read with 26pt and the rest of dyslexic participants preferred 22pt. They reported that none of the participants preferred small (14pt) or medium (18pt) for reading. Since the present study had no significant difference between small and medium font size, it is unknown if large (22pt) and extra large (26pt) font size will have different effects on the preferences.

According to study by Rello et al. (2013), reading with larger font size will significantly increase comprehension scores as shown in Figure 2.12. Thus, reading with medium font size lead to better comprehension scores than small font size. Contrary to the present study, no correlation was found between font size and comprehension score in this study as found by others in studies involving university students (Chandler, 2001; Tavakoli & Kheirzadeh, 2011). As Rello et al.'s (2013) study had a mixture of children and adults in their participant group, their results might have different significant effects compared to proficient readers who require smaller font sizes to enhance reading comprehension (Katzir, Hershko, & Halamish, 2013).

3.3.4.3 Overall Effects of Participant Group

Regarding the differences between the groups, no previous study on the effects of typeface and font size has compared participants without dyslexia, participants with mild dyslexia and participants with moderate dyslexia. Therefore, no comparison between these three groups could be made with other studies. However, a review of 20 years of eye tracking research with dyslexia showed various inconsistent results of their eye movements (Rayner, 1998).

It is widely accepted that poor readers and people with dyslexia are different from normal readers in that they tend to make more and longer fixations, shorter saccades and more regressions (Rayner, 1998). Supporting that, this study found that moderate dyslexic participants made a significantly higher number of fixations compared to mild dyslexic and non-dyslexic participants. However, there was no significant difference between the groups in the duration of fixations. However, there was an interesting interaction between Group and Typeface on fixation duration, with the moderate dyslexic and non-dyslexic participants having significantly shorter fixation durations on the dyslexia optimised typefaces compared to either serif or sans serif fonts. This was one of the few results favouring the dyslexia-optimised typefaces (the other being a main effect for typeface on fixation duration, discussed below).

As discussed by previous research, people with dyslexia may have problems in reading comprehension (Snowling, 2012; Vellutino et al., 2004). However, this present study found no significant relationship between participants groups and their reading comprehension. This result shows that participants with dyslexia in this study can read relatively well and understand what they have read. As discussed by Snowling (1998), many people with dyslexia have typical comprehension skills despite their decoding difficulties.

3.3.4.4 Limitations of This Study

This study had a difficulty in finding participants in the middle of the University term especially with dyslexic participants. I found more mild dyslexic participants compared to moderate dyslexics. Some of the moderate dyslexics were reluctant to participate after being told that the tasks would involve lots of reading. In term of materials prepared for the study, some participants had little interest about the Olympics theme and easily bored reading the texts. In addition, some participants found reading the texts was quite long and exhausting, especially when they had the experimental session in the evening. As majority of participants were university student, dyslexic participants in this study might be a good reader compared to dyslexic without higher education. Dyslexic participants in this study might have compensation strategies in reading.

In the data preparation, I realised some participants had an erratic eye gaze due to the natural lighting coming from the outside. The quiet room used in this study has two medium sized windows and during the session, the blinds of the windows were not properly covering the window frames. Due to sudden movement of cloud during windy days, the angle of lighting coming from the windows differed during the session and affected the eye gaze tracking. On account of the difficulty in finding participants, I had decided to measure the number of fixations, fixation duration and reading time based on number of words per paragraph, to be able to use those which had accurate eye gaze. Ideally, all data for problematic eye gaze would to be fully replaced as it added extra complexity during data preparation.

In term of experimental design, the texts were not counterbalanced across all typeface and font size combinations. However, all texts were measured to have similar length, number of paragraphs, complex words and readability scores (tested with FKRE, FKGL and GFS), the results possibly influenced by the language variables.

3.3.4.5 Conclusions

Based on the discussion above, it can be seen that some of the objective data (from eye gaze) and subjective data (from post-study questionnaire) appear to contradict each other. From this study, it can be seen that participants had fewer fixations with small font size and shorter fixation duration with dyslexia-optimised typeface in the objective measures. For subjective measures, participants preferred sans serif typeface, and had different levels of comfort with dyslexia-optimised typefaces. Dyslexic participants found OpenDyslexic typeface was not easier than standard typefaces and they were not interested to use the typeface.

In addition, this study also demonstrates the variability of results in the research on the effects of typeface and font size on reading for dyslexic and non-dyslexic participants.

However, I am confident about the results of this study since all the significant effects found had very large effect sizes. Large effect size quantifies robust difference between groups and reduces the likelihood of an accident of sampling (Coe, 2002).

CHAPTER 4: STUDY 2 – EFFECTS OF LINE SPACING AND LINE LENGTH ON DETAILED READING

4.1 Introduction

This chapter describes Study 2 in this programme of research which aims to answer Research Questions 1(c) and 1(d) discussed in Chapter 1:

1 (c) To what extent does the line spacing used in the presentation of text on a computer screen affect eye gaze behaviour, reading performance, comprehension and preferences of adults with dyslexia compared to adults without dyslexia?

1 (d) To what extent does the line length used in the presentation of text on a computer screen affect eye gaze behaviour, reading performance, comprehension and preferences of adults with dyslexia compared to adults without dyslexia?

As was seen from the discussion in Chapter 2, a number of studies have investigated how line spacing and line length might affect people with dyslexia (Rello & Baeza-Yates, 2015; Rello et al., 2012b; Rello, Pielot, et al., 2013; Schneps, Thomson, Sonnert, et al., 2013; Schoonewelle, 2013; Zorzi et al., 2012). However as discussed, there is no strong evidence on which particular line spacing and line length can make reading easiest for people with dyslexia.

Similar to Study 1, presented in Chapter 3, this study will distinguish people with dyslexia into two levels based on their severity of dyslexia. A well-established diagnostic tool will be used to identify participants as mild and moderate dyslexia. It is interesting to understand if the severity of their dyslexia plays an important role on the effects of the typography aspects, namely line spacing and line length, on their reading.

As reading comprehension is measured in this study, a preparatory study was therefore conducted to test the accuracy and difficulty of comprehension questions to be used. This chapter will present the method and results of the preparatory study. Materials prepared in the preparatory study will be used in the Study 2. The rest of this chapter also will describe methods, results and discussion for Study 2.

4.2 Preparatory Study – Assessment of Question Difficulty

This preparatory study aimed to test the accuracy and difficulty of each comprehension question which would be used in the Study 2. Four texts were developed to be used in the Study 2 and their characteristics are discussed in Section 4.3.1.3. Multiple choice questions for each text were developed based on findings from Davey (1988), where correct answer for each question need to be stated explicitly in the text.

Initially, participants were asked to read texts and answer proposed multiple choice questions for each of the texts for Study 2. They needed to answer and rate on the questions' difficulty twice, without referring to the text and while referring back to the text. They also needed to state their prior knowledge about the questions. This preparatory study enabled me to make a selection of which questions would be used in Study 2.

4.2.1 Method

Method for this study employed similar method in the preparatory study for Study 1 as discussed in Section 3.2.1.

4.2.1.1 Design

A repeated measures design was chosen to investigate the difficulties of the questions, how accurately participants could answer the questions and whether they already knew the answer for the questions before reading the texts. Two types of difficulties were measures: Question Difficulty Rating (without referring to the text) and Answer Difficulty Rating (while referring back to the text). Questions accuracy was measured by participants' correct answers to the multiple choice questions. Previous Knowledge Assessment was used to measure participants' familiarity by asking whether they already knew the answer before reading the text. Overall, this preparatory study employed same design as discussed in previous study, Section 3.2.1.1.

4.2.1.2 Participants

Participants were all English native speakers. Five participants (4 males and 1 female) were participated. Their ages ranged from 27 to 66 years old.

4.2.1.3 Materials

Four texts and questions were developed which would be used for four experimental tasks in Study 2. Two texts had a theme about fruit and vegetable, while another two had a theme about cities. All texts were adapted from the Wikipedia[11,12,13,14] website. The characteristics of the texts are described in Section 4.3.1.3 of Study 2.

The comprehension questions for each text were four multiple choice questions with three response alternatives. The comprehension questions were developed based on work by Davey (1988), where the correct answer needs to be stated explicitly in one of the passage in the text. When possible, one question related to each paragraph in the text, but the questions were presented in a random order regardless of which paragraph they came from. Figure 4.1 shows one of the multiple choice questions which had one correct answer and two alternative distractor choices.

A predictable quality and quantity of avocados is achieved by …

a. growing the trees in humid conditions
b. propagating the trees by grafting
c. controlling the amount of fertilisers for the trees

Figure 4.1 Sample of multiple choice question and answers

Each of the comprehension questions was accompanied by two difficulty ratings and a familiarity question. Both difficulty ratings were rated on 9-point Likert items. Participants' familiarity on the subject matter of the questions was measured by asking whether they already knew the answer before reading the text using Yes/No option. Appendix K shows the four sets of texts and a sample of one set of comprehension questions with its accompanying difficulty ratings and familiarity question.

4.2.1.4 Procedure

Participants were emailed a file of the texts and questions. They were asked to read the texts and answer the accompanying questions. After they finished reading the text, they first answered the multiple choice question without looking at the text and then rated the difficulty of the question on a 9-point Likert scale item (the Question Difficulty

[11] https://en.wikipedia.org/wiki/Avocado
[12] https://en.wikipedia.org/wiki/Lettuce
[13] https://en.wikipedia.org/wiki/Chennai
[14] https://en.wikipedia.org/wiki/Kyoto

Rating). They then looked back at the text, rated again the difficulty of each question on a 9-point Likert scale item (the Answer Difficulty Rating) and indicated whether they were familiar with the subject matter and already knew the answer for that question before reading the text (the Previous Knowledge Assessment). After they had finished all the questions and difficulty assessments, they returned the file through email or on paper.

4.2.2 Results and Discussion

For each question in the text, participants' accuracy was averaged into Correct Answer, participants' difficulty rating without referring to the text were averaged into Question Difficulty Rating, participants' difficulty rating while referring back to the text were averaged into Answer Difficulty Rating and participants' familiarity on the subject matter were averaged into Previous Knowledge Assessment.

A total of three questions was needed for each text for Study 2. Various factors contributed to the elimination of particular questions. For each text, a question was discarded if it was too difficult, too easy or most participants have already knew the answer before reading the text. The aim was to have questions with similar difficulty ratings as those in Study 1 (see Section 3.2.2). In that study, Question Difficulty Rating were in the range of 4.00 to 7.75 (on a scale 1 to 9), Answer Difficulty Rating were in the range 5.00 to 8.75 (on a scale 1 to 9) and Previous Knowledge Assessment scores were in the range 0.00 to 0.50 (on a scale 0 to 1).

Table 4.1 shows the results for all texts in the current preparatory study. For each text, question marked in red indicates it was discarded from being used Study 2. Means for all questions that have been selected for each text was also shown in the table.

From Table 4.1, Experimental Text 1 (M = 1.00) had the most accurate questions (correct answers from all participants), while Experimental Text 2 (M = 0.67) had the lowest score for accuracy. Without looking back at the text, Experimental Text 2 (M = 5.90) had the most difficult questions, while Experimental Text 1 had the easiest questions (M = 7.53). While looking back at the text, participants had the most difficulty in finding answer for Experimental Text 4 (M = 8.07), and Experimental Text 1 (M = 8.33) had the easiest rating. Participants were not familiar with questions in Experimental Text 2 (M = 0.00).

Table 4.1 Results for all texts in the preparatory study

Text	Question	Correct Answer (*M*)	Questions Difficulty Rating (*M*)	Answer Difficulty Rating (*M*)	Previous Knowledge Assessment (*M*)
Experimental 1	1	1.00	8.20	8.60	0.60
	2	1.00	7.80	8.20	0.40
	3	1.00	7.80	8.40	0.40
	4	1.00	7.00	8.40	0.00
	Selected questions (*M*)	1.00	7.53	8.33	0.27
Experimental 2	1	0.60	5.80	8.40	0.00
	2	0.80	5.90	8.20	0.00
	3	1.00	5.40	8.40	0.00
	4	0.60	6.00	7.70	0.00
	Selected questions (*M*)	0.67	5.90	8.10	0.00
Experimental 3	1	1.00	5.80	7.90	0.00
	2	1.00	6.60	8.40	0.00
	3	0.80	6.70	7.90	0.00
	4	1.00	8.40	8.60	0.40
	Selected questions (*M*)	0.93	7.23	8.30	0.13
Experimental 4	1	1.00	7.80	8.40	0.20
	2	0.40	5.00	7.90	0.00
	3	0.80	7.60	8.10	0.00
	4	1.00	6.10	7.70	0.00
	Selected questions (*M*)	0.93	7.17	8.07	0.07

Note. Questions marked in red were discarded.

The preparatory study allowed me to create sets of comprehension questions with similar levels of difficulty, which would not be familiar to participants. The means for all texts for their Correct Answer were 0.90±0.16, Question Difficulty Rating were 6.92±0.74, Answer Difficulty Rating were 8.23±0.19 and Previous Knowledge Assessment 0.13±0.14. These questions and answers will now be used in the Study 2. As was discussed in Chapter 2, previous studies only measured difficulty of text passages read by participants and none of the studies measured difficulty of the questions which were asked after reading to measure reading comprehension (Rello & Baeza-Yates, 2015; Rello et al., 2012b; Rello, Pielot, et al., 2013; Schneps, Thomson, Sonnert, et al., 2013; Zorzi et al., 2012).

4.3 Study 2

As discussed in Chapter 2, small increments in spacing can have an impact on crowding effects from neighbouring letters (Levi, 2008). Crowding is said to influence the performance of reading in poor readers and people with dyslexia (Levi, 2008; Martelli et al., 2009). Several studies have investigated how character spacing and word spacing can have an impact on reading with people with dyslexia (Martelli et al., 2009; Moll & Jones, 2013; Perea et al., 2012; Sjoblom et al., 2016; Spinelli et al., 2002; Zorzi et al., 2012). Little attention has been given to the vertical crowding effect, particularly the effects of line spacing on reading with people with dyslexia (Bernard et al., 2007; Rello & Baeza-Yates, 2015; Rello et al., 2012b; Rello, Pielot, et al., 2013; Schoonewelle, 2013). Furthermore, I found limited numbers of studies on line length for people with dyslexia (Rello & Baeza-Yates, 2015; Rello et al., 2012b; Schneps, Thomson, Sonnert, et al., 2013). It is important to identify line length for reading since it will make the return sweep to the beginning of the next line harder if it is too long (Dyson, 2004; Dyson & Haselgrove, 2001). Scarcity of research on these two areas with participants with dyslexia stressed the need for this study.

4.3.1 Method

Method for this study employed similar method in Study 1 as discussed in Section 3.3.1.

4.3.1.1 Design

This study is a 3 × 2 × 2 mixed-participants design with participant groups as the between participants variable and line spacing and line length as the repeated measures variables. Participants were divided into three groups (non-dyslexic, mild dyslexic and moderate dyslexic) on the basis of a diagnostic checklist for dyslexia (Snowling et al., 2012). For the line spacing variable, two levels of line spacing were used: 1.5 (or 150% of 18-point size, 27pt) and 2.0 (or 200% of 18-point size, 36pt). For the line length variable, two levels of length were used: 60 – 70 characters per line (cpl) and 80 – 90 cpl. The line length defined in this study counts all characters including spaces in one line.

The dependent variables were measured into three groups of measurements: (1) eye gaze behaviour, (2) reading performance, and (3) participants' preferences.

(1) The eye gaze behaviour measures are: number of fixations and average fixation duration.

(2) The reading performance measures are: reading time and comprehension scores. Comprehension scores were only measured to make sure participants read the text.
(3) The participants' preferences measures are: rating of ease of reading and preferred combination of line spacing and line length. All ratings were gathered in a post-study questionnaire on 7-point Likert scale items.

There were four combinations of line spacing and line length. Participants read four texts, each one with a different combination. The order of the texts and the version were counterbalanced.

Ethics approval for this study has been granted by Physical Sciences Ethics Committee of the University of York. Participants were ensured about the confidentiality and anonymity of their data.

4.3.1.2 Participants

Participants were asked to complete an online screening diagnostic tool. They were categorised into non-dyslexic group if their diagnostic scores ranged between 0 to 10, mild dyslexic group if their scores ranged from 11 to 21 and moderate dyslexic group from 22 to 32. A total of 24 participants took part in this study, 8 participants in each group (non-dyslexic, mild dyslexic and moderate dyslexic). All participants were native English speakers with normal vision or vision correctable with eyewear.

The non-dyslexic group comprised 4 females and 4 males with ages ranging from 18 to 25 years (M = 21.00, SD = 2.62). Five were undergraduate students, one was a Masters student (and also working part-time) and two were PhD students. All of them had more than 6 years of Web experience. Mean of Web expertise is 5.63 (SD = 1.06), on a 7-Likert item.

The mild dyslexic group comprised 4 females and 4 males with ages ranging from 19 to 30 years (M = 21.88, SD = 4.22). Six were undergraduate students, one was PhD student and one was employed. Two of them had 4 – 6 years of Web experience and the rest had more than 6 years of Web experience. Mean of Web expertise is 6.00 (SD = 0.53), on a 7-Likert item.

The moderate dyslexic group comprised 4 females and 4 males with ages ranging from 18 to 20 years (M = 19.25, SD = 1.13). All of them were undergraduate students. One of them had 4 – 6 years of Web experience and the rest had more than 6 years of Web experience. Mean of Web expertise is 5.75 ($SD = 0.46$), on a 7-Likert item.

Participants were given a £10 gift e-voucher from Amazon to thank them for participating in the study.

4.3.1.3 Materials and Equipment

All equipment (laptop, computer screen, eye tracker camera, cables, wireless keyboard and mouse) and dual monitor setup were the same as in Study 1 (see Section 3.3.1.3 and Figure 3.4). Participants sat on a fixed chair during this study. Similar viewing distance of 60cm applied between participants and the monitor screen.

An online ARQ (Snowling et al., 2012), informed consent form and online pre-study questionnaire used in this study were the same as in Study 1 (refer Section 3.3.1.3).

Four sets of experimental texts and comprehension questions (consisting of three multiple choice questions for each text) were created. The questions were developed and tested in the preparatory study describes in Section 4.2. Two sets for practice texts and comprehension questions were adopted from (Chatrangsan, In Preparation). The practice texts had a theme about fruit and places of interest. One experimental text had a theme about fruit, one about a vegetable and another two had a theme about cities. All experimental texts were adapted from Wikipedia[15,16,17,18]. In each text, the number of sentences, percentage of sentences with more than 20 syllables, percentage of words more than 12 letters and number of sentences with passive voice were controlled within ± 12% range from the mean value. In order to control these scores to between ± 12% range from the mean value, texts were repeatedly reworded, replaced and restructured. Texts created were measured with the premium version of an online readability testing tool, readable.io[19]. The texts can be found in Appendix K.

---

[15] https://en.wikipedia.org/wiki/Avocado
[16] https://en.wikipedia.org/wiki/Lettuce
[17] https://en.wikipedia.org/wiki/Chennai
[18] https://en.wikipedia.org/wiki/Kyoto
[19] https://readable.io

Practice texts had similar characteristics with the experimental texts however the practice texts had lower scores on readability measures than the experimental texts, which means the practice texts were easier. Each text has the following characteristics:

- Between 225 to 235 words
- Between 13 to 16 sentences
- 3 paragraphs
- 0% of words had more than 12 letters
- Between 31% t0 44% of sentences had more than 20 syllables
- Between 3 to 5 sentences with passive voice

To ensure all the experimental texts were comparable, a number of readability measures were controlled within ± 12% of the mean value. The readability measures used comprised the FKRE, FKGL, GFS and average grade level. FKRE score range from 0.0 – 100.0, FKGL score and GFS score equivalent to US grade level from 0.0 – 20.0 (Armstrong, 1980; Colmer, 2019a, 2019b; Cotugna et al., 2005; Flesch, 1948; Kincaid et al., 1975; Klare, 1984; Wilson & Corlett, 2005). Compared to readability measures discussed in Section 3.3.1.3, a new readability measure was added, average grade level. It represents an average of grade levels calculated using variety of readability formulas such as FKGL, GFS, Coleman-Liau Index, Simple Measure of Gobbledygook (SMOG) Index and Automated Readability Index. Readability scores for the experimental texts and their characteristics are shown in Table 4.2. Scores for the readability measures for the experimental texts are described below:

- FKRE
    - Score between 43.5 to 55.5, indicates the text is difficult and represents general academic-oriented articles
    - Suitable for senior high school students and average readers in first year university
- FKGL
    - Score between 9.1 to 11.6, indicates the text as suitable for US 9th grade to 12th grade and equivalent to UK Year 10 to Year 13
    - Easily understood by average readers in first year university
- GFS
    - Score between 9.9 to 12.7, represents leading magazine articles

    - Suitable for all high school students to first year university students
- Grade level
    - Score between 9.8 to 12.0, indicates the text as suitable for US 10th grade to 12th grade and equivalent to UK Year 11 to Year 13
    - Suitable for all senior high school students

Table 4.2 Results of readability scores for experimental texts used in Study 2

Text	No. of Words	No. of Sentences	FKRE	FKGL	GFS	Avg. Grade Level	Sentences > 20 syllables (%)	Words > 12 letters (%)	Passive Voice Sentences (%)
1	233	16	47.5	10.2	11.1	10.8	44	0	3
2	230	15	50.2	10.0	10.5	10.3	40	0	4
3	232	13	47.6	11.1	11.6	11.6	31	0	3
4	235	14	52.6	10.1	12.2	11.0	36	0	5
M	**232.50**	**14.50**	**49.48**	**10.35**	**11.35**	**10.93**	**37.75**	**0.00**	**3.75**
SD	**2.08**	**1.29**	**2.43**	**0.51**	**0.72**	**0.54**	**5.56**	**0.00**	**0.96**

Note. 'FKRE' denotes Flesch-Kincaid Reading Ease; 'FKGL' denotes Flesch-Kincaid Grade Level; 'GFS' denotes Gunning Fog Score.

The experimental texts and comprehension questions were presented in four different versions. Each version had a different combinations of the line spacing and line length. Practice texts used two other versions (line spacing of 1.0 with 70 – 80cpl, and line spacing of 1.0 with 90 – 100cpl) which were not included in any of the experimental texts. Texts were presented in the middle of the participants' screen, left-justified. All texts were presented in Arial typeface, 18-point size with black text on a white background. Figure 4.2 – Figure 4.5 illustrate one of the experimental texts in each of the versions when displayed on the participant's monitor screen. Comprehension questions used the same version of line spacing-line length combination as the corresponding text.

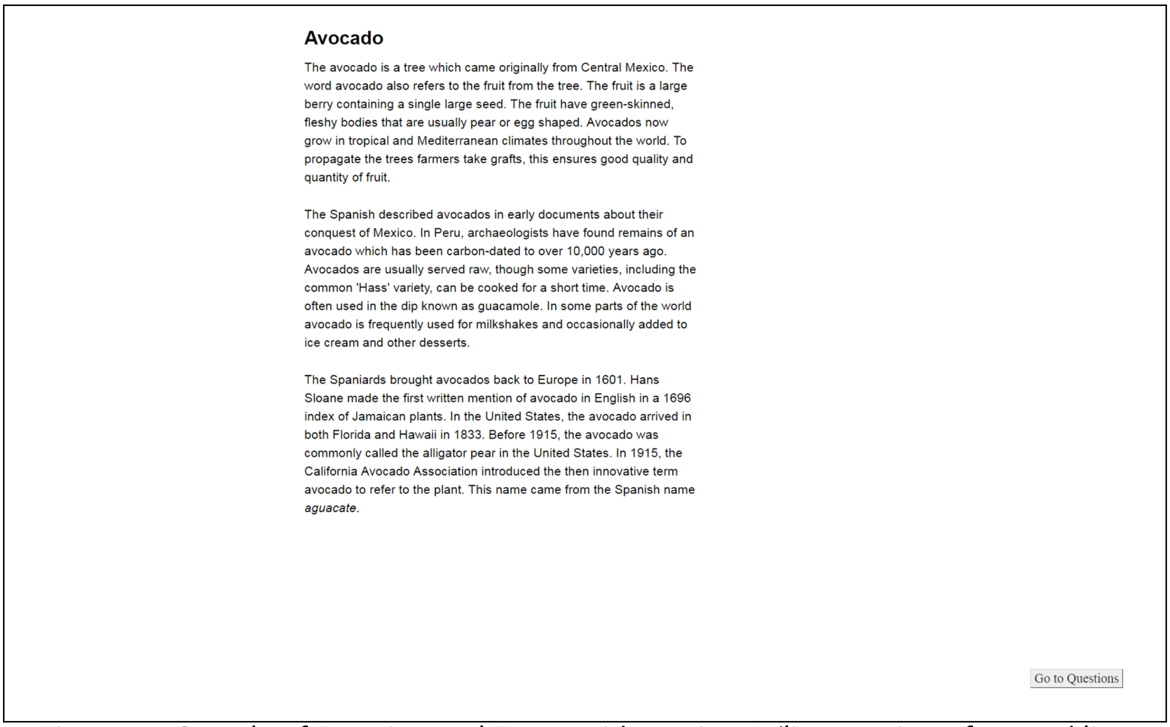


Figure 4.2 Sample of Experimental Text 1 with version A (line spacing of 1.5 and line length of 60 – 70cpl) used in Study 2

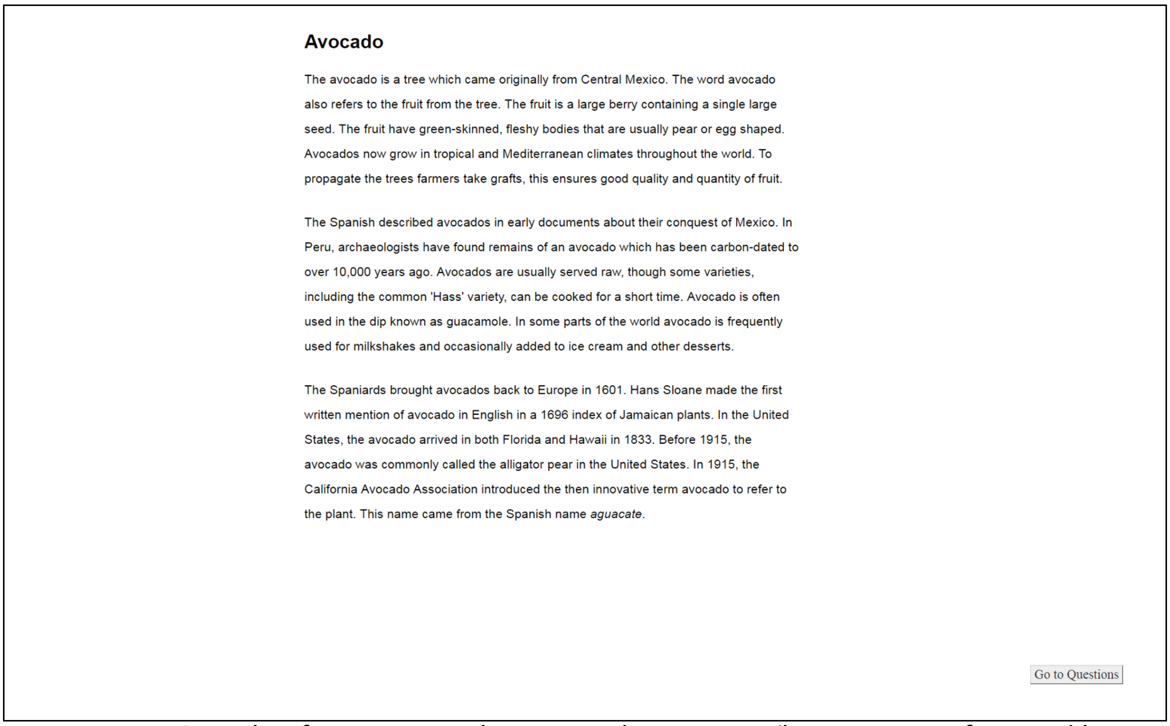
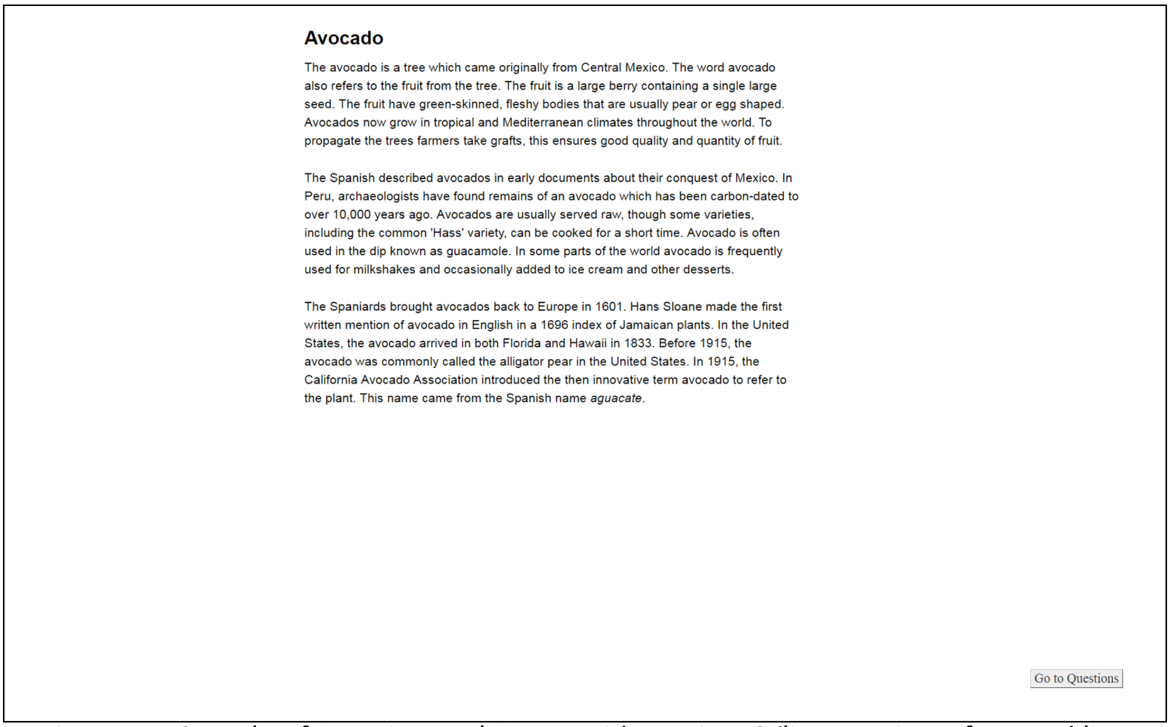


Figure 4.3 Sample of Experimental Text 1 with version B (line spacing of 2.0 and line length of 80 – 90cpl) used in Study 2



Avocado

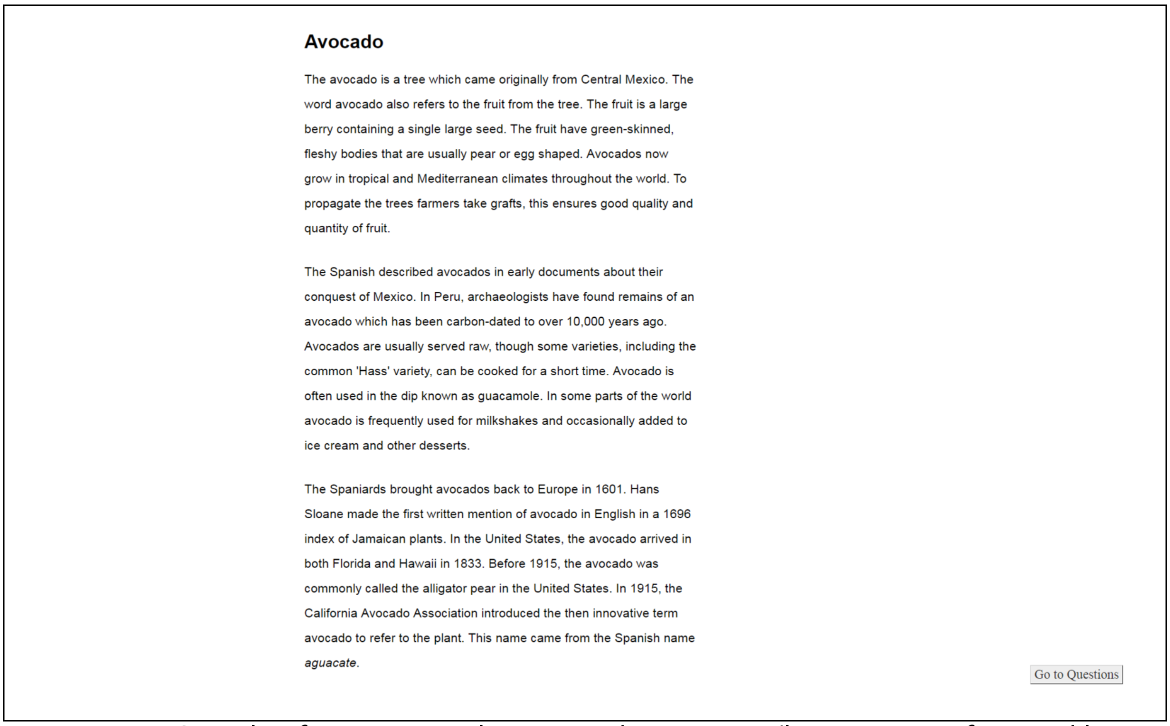
The avocado is a tree which came originally from Central Mexico. The word avocado
also refers to the fruit from the tree. The fruit is a large berry containing a single large
seed. The fruit have green-skinned, fleshy bodies that are usually pear or egg shaped.
Avocados now grow in tropical and Mediterranean climates throughout the world. To
propagate the trees farmers take grafts, this ensures good quality and quantity of fruit.

The Spanish described avocados in early documents about their conquest of Mexico. In
Peru, archaeologists have found remains of an avocado which has been carbon-dated to
over 10,000 years ago. Avocados are usually served raw, though some varieties,
including the common 'Hass' variety, can be cooked for a short time. Avocado is often
used in the dip known as guacamole. In some parts of the world avocado is frequently
used for milkshakes and occasionally added to ice cream and other desserts.

The Spaniards brought avocados back to Europe in 1601. Hans Sloane made the first
written mention of avocado in English in a 1696 index of Jamaican plants. In the United
States, the avocado arrived in both Florida and Hawaii in 1833. Before 1915, the
avocado was commonly called the alligator pear in the United States. In 1915, the
California Avocado Association introduced the then innovative term avocado to refer to
the plant. This name came from the Spanish name *aguacate*.

Go to Questions

Figure 4.4 Sample of Experimental Text 1 with version C (line spacing of 1.5 and line length of 80 – 90cpl) used in Study 2



Avocado

The avocado is a tree which came originally from Central Mexico. The
word avocado also refers to the fruit from the tree. The fruit is a large
berry containing a single large seed. The fruit have green-skinned,
fleshy bodies that are usually pear or egg shaped. Avocados now
grow in tropical and Mediterranean climates throughout the world. To
propagate the trees farmers take grafts, this ensures good quality and
quantity of fruit.

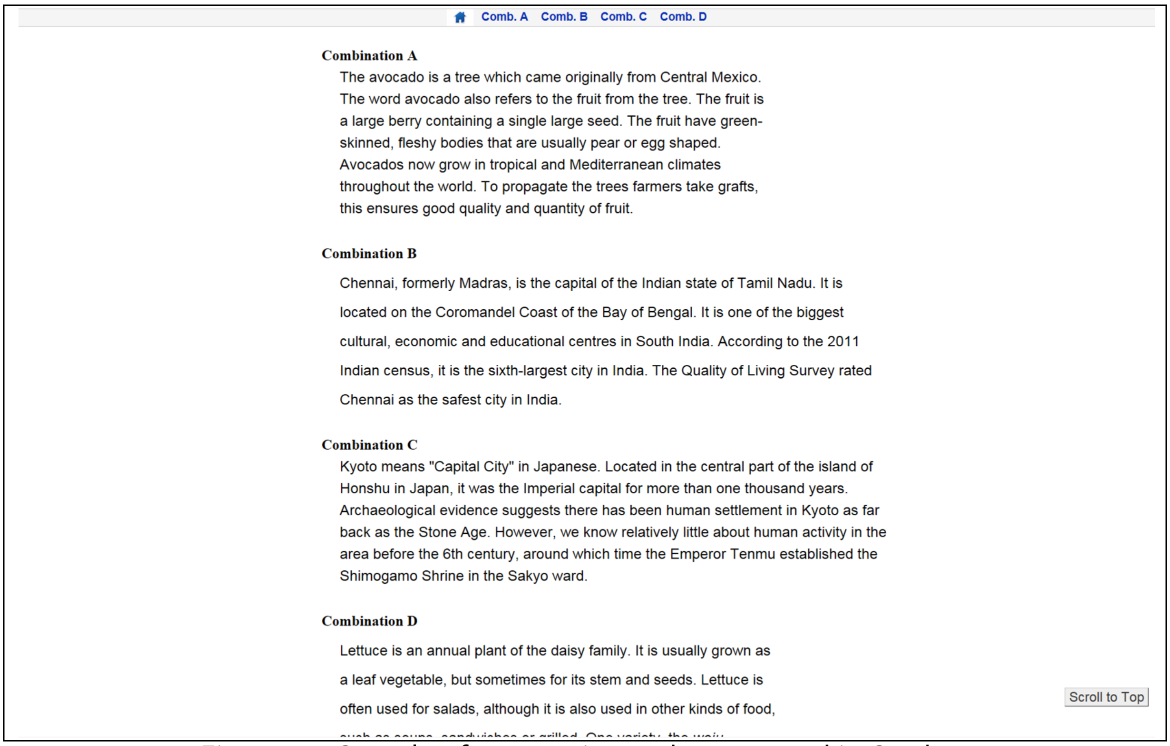
The Spanish described avocados in early documents about their
conquest of Mexico. In Peru, archaeologists have found remains of an
avocado which has been carbon-dated to over 10,000 years ago.
Avocados are usually served raw, though some varieties, including the
common 'Hass' variety, can be cooked for a short time. Avocado is
often used in the dip known as guacamole. In some parts of the world
avocado is frequently used for milkshakes and occasionally added to
ice cream and other desserts.

The Spaniards brought avocados back to Europe in 1601. Hans
Sloane made the first written mention of avocado in English in a 1696
index of Jamaican plants. In the United States, the avocado arrived in
both Florida and Hawaii in 1833. Before 1915, the avocado was
commonly called the alligator pear in the United States. In 1915, the
California Avocado Association introduced the then innovative term
avocado to refer to the plant. This name came from the Spanish name
aguacate.

Go to Questions

Figure 4.5 Sample of Experimental Text 1 with version D (line spacing of 2.0 and line length of 60 – 70cpl) used in Study 2

An online post-study questionnaire (see Appendix L) was developed to measure participants' ratings of the ease of reading and preferences for the combinations of line spacing and line length. In the questionnaire, participants were asked to rate ease of reading for each combination, then chose one combination of line spacing and line length they most preferred. In order for participants to rate the ease of reading and give their overall preference, a web page was developed which displayed excerpts of experimental texts they had read with the different versions of the texts. This companion web page assisted participants in answering the questionnaire easily. Figure 4.6 shows the companion web page for the questionnaire with all the versions.


Comb. A Comb. B Comb. C Comb. D

Combination A

The avocado is a tree which came originally from Central Mexico. The word avocado also refers to the fruit from the tree. The fruit is a large berry containing a single large seed. The fruit have green-skinned, fleshy bodies that are usually pear or egg shaped. Avocados now grow in tropical and Mediterranean climates throughout the world. To propagate the trees farmers take grafts, this ensures good quality and quantity of fruit.

Combination B

Chennai, formerly Madras, is the capital of the Indian state of Tamil Nadu. It is located on the Coromandel Coast of the Bay of Bengal. It is one of the biggest cultural, economic and educational centres in South India. According to the 2011 Indian census, it is the sixth-largest city in India. The Quality of Living Survey rated Chennai as the safest city in India.

Combination C

Kyoto means "Capital City" in Japanese. Located in the central part of the island of Honshu in Japan, it was the Imperial capital for more than one thousand years. Archaeological evidence suggests there has been human settlement in Kyoto as far back as the Stone Age. However, we know relatively little about human activity in the area before the 6th century, around which time the Emperor Tenmu established the Shimogamo Shrine in the Sakyo ward.

Combination D

Lettuce is an annual plant of the daisy family. It is usually grown as a leaf vegetable, but sometimes for its stem and seeds. Lettuce is often used for salads, although it is also used in other kinds of food,

Scroll to Top

Figure 4.6 Sample of companion web page used in Study 2

4.3.1.4 Procedure

The overall procedure was quite similar to Study 1 (refer Section 3.3.1.4). Before the experiment session, participants were emailed a hyperlink to the online ARQ. Participants completed the online ARQ (see Appendix C) to screen them into non-dyslexic, mild dyslexic and moderate dyslexic groups. During the experiment session, they were asked to complete an informed consent form (see Section A in Appendix D) and pre-study questionnaire about their demographic information and their use of the Web (see Appendix E).

In the experimental session, the participant was asked to sit comfortably on a fixed chair with a distance of approximately 60cm to computer screen. The participant started the experiment with calibration in which they needed to focus their eye gaze on five red dots on the screen. In the event of unsuccessful calibration, the process was repeated. After a successful calibration and validation, they completed two practice tasks. In each task, the participant read one text and answered the corresponding set of comprehension questions about the text.

Then, they repeated the calibration and validation process before starting the four experimental tasks. In each task, they read the text and clicked on a 'Go to Questions' button after finished reading the text. They were then presented with a set of questions about the text they had read and answered the questions. Appendix J shows the sequence of the texts which participants read including how the sequence was made.

After each participant finished all four experimental tasks, they completed the online post-study questionnaire (see Appendix L). The participant was shown a web page which presented all the excerpts from all the experimental texts with the different combination of line spacing and line length. Then, participants were debriefed and any questions about the study were answered. Participants signed Section B in the Consent Form (see Appendix D) and were asked to give their email address for the e-voucher. They received the gift e-voucher through email from Amazon. Finally, they were asked whether they would like to see their example of their eye gaze pattern before they left the room.

4.3.2 Data Preparation

Results from this study were prepared in a similar manner as discussed in Section 3.3.2. Using the SMI BeGaze™, an area of interest (AOI) was created for the whole text. Results for the number of fixations, fixation duration and text reading time were taken from data recorded with the SMI Experiment Center™. Comprehension scores were taken from a MySQL database which was recorded during participants' interaction with the web pages. Their preferences were taken from the post-study questionnaire given after the experimental task.

The number of fixations, fixation duration and reading time was calculated from the time when the participant's eye gaze entered the AOI in the text to the time when the participant pressed the button to move to the comprehension questions web page.

Fixations outside the AOI were not counted. Average fixation duration was calculated by dividing the fixation duration and number of fixations in AOIs.

The results were grouped into three categories: (1) eye gaze behaviour (number of fixations and average fixation duration), (2) reading performance (text reading time and comprehension score) and (3) participants' preferences.

Each text read by participants was visually inspected using the SMI BeGaze^{TM} to make sure all eye gaze data were accurate. One initial participant had inaccurate eye gaze recorded outside AOI. The participant was replaced with another participant matched for age, gender and their severity of dyslexia.

All data were entered into a MS Excel spreadsheet and then analysed using the SPSS statistical program. All data were visually analysed using histograms to check for normal distribution. Outliers for number of fixations, average fixation duration and text reading time were adjusted using the *winsorization* technique (DeCator, 2015). In this technique, outliers were adjusted if the values below or above than $Mean \pm 2SD$.

4.3.3 Results

The majority of data for reading performance, eye gaze behaviour and preferences were normally distributed. However, in less than half of the data for eye gaze behaviour and reading performance were the variances heterogenous. According to Glass et al. (1972), non-normality and heterogenous data can additively affect the results of ANOVA. In order to have a valid result for this study, the data was transformed using Log10 in an attempt to reduce the non-normality and heterogenous data. After the transformation, I found no changes to the normality of data and the homogeneity of variances. I decided to run both parametric (with three-way ANOVA) and non-parametric tests on the original data. However for non-parametric tests, there is no test equivalent to the three-way ANOVA (Laerd Statistics, 2015). Therefore, the ANOVA was then conducted on both original data and transformed data. Significant results with both tests were then compared.

For both original and transformed data, similar significance results were found with the ANOVA. I decided to report the results of this study using the original data with the parametric test of ANOVA since the ANOVA is robust enough to handle the violations when the sample size is equal (Blanca et al., 2017; Laerd Statistics, 2015; Schmider et al., 2010).

A 3 x 2 x 2 mixed ANOVA was used, comprising three participant groups (non-dyslexic, mild dyslexic and moderate dyslexic), two line spacings (1.5 and 2.0) and two line lengths (60 – 70cpl and 80 – 90cpl). The dependent variables were number of fixations, average fixation duration, text reading time and participants' preferences.

Since comprehension scores were used to confirm that participants had read the text, relationships between the scores and the three participant groups, two line spacings and two line lengths were investigated. For total comprehension scores, a parametric tests one-way ANOVA was conducted, comprising three participant groups (non-dyslexic, mild dyslexic and moderate dyslexic) with total comprehension scores as a dependent variable. Comprehension scores for the two line spacings and two line length were not normally distributed, therefore non-parametric tests were used.

Bonferroni-corrected post-hoc tests were conducted to investigate specific differences uncovered by the analyses. An alpha significance level of .05 was used in all statistical tests in this study. For effect size, partial eta squared of .01 is small, .06 is medium and .14 is large was used to determine the magnitude of the statistical results (Draper, 2018). Details for each test are discussed in next sections.

4.3.3.1 Eye Gaze Behaviour

a. Number of Fixations

Table 4.3 shows the descriptive statistics for number of fixations for the three participant groups. The number of fixations were normally distributed as assessed by Shapiro-Wilk's test of normality. The assumption of homogeneity of variances was met except for Line Spacing 1.5 with Line Length 60 – 70cpl (F = 4.95, p = .02) and Line Spacing 2.0 with Line Length 80 – 90cpl (F = 4.71, p = .02), as assessed by Levene's test for equality of variances. Mauchly's test of sphericity indicated that the assumption of sphericity was met. The results of the ANOVA are summarised in Table 4.4.

Table 4.3 Descriptive statistics of number of fixations for all participant groups

	Non-Dyslexic		**Mild Dyslexic**		**Moderate Dyslexic**	
	M	***SD***	***M***	***SD***	***M***	***SD***
LS of 1.5, LL of 60 – 70cpl	262.75	70.98	282.38	70.09	376.88	136.08
LS of 1.5, LL of 80 – 90cpl	244.63	63.92	270.13	54.82	333.13	87.66
LS of 2.0, LL of 60 – 70cpl	243.63	56.63	295.63	79.76	328.38	99.90
LS of 2.0, LL of 80 – 90cpl	245.50	48.74	274.88	60.91	332.63	100.50

Note. 'LS' denotes Line Spacing; 'LL' denotes Line Length. Measures count per text.

Table 4.4 Summary of three-way mixed ANOVA on number of fixations

Effects		***F***	***p***	η_p^2
Main	Group	*F*(2, 21) = 3.11	ns	.23
	Line Spacing	*F*(1, 21) = 1.40	ns	.06
	Line Length	*F*(1, 21) = 5.93	.02	.22
Two-way interaction	Group × Line Spacing	*F*(2, 21) = 1.95	ns	.16
	Group × Line Length	*F*(2, 21) = 0.33	ns	.03
	Line Spacing × Line Length	*F*(1, 21) = 2.54	ns	.11
Three-way interaction	Group × Line Spacing × Line Length	*F*(2, 21) = 1.72	ns	.14

Note. 'ns' indicates not statistically significant.

There was a significant main effect for Line Length with large effect size, *F*(1, 21) = 5.93, *p* = .02, η_p^2 = .22 where 60 – 70cpl (*M* = 298.27, *SD* =95.52) had significantly higher number of fixations than 80 – 90cpl (*M* = 283.48 , *SD* = 77.38).

b. Average Fixation Duration

Table 4.5 shows the descriptive statistics for the average fixation duration for the three participant groups. Average fixation duration were normally distributed as assessed by Shapiro-Wilk's test of normality. The assumption of homogeneity of variances was met except for Line Spacing 1.5 with Line Length 80 – 90cpl (*F* = 4.06, *p* = .03), as assessed by Levene's test for equality of variances. Mauchly's test of sphericity indicated that the assumption of sphericity was met. The results of the ANOVA are summarised in Table 4.6.

Table 4.5 Descriptive statistics for average fixation duration (ms) for all participant groups

	Non-Dyslexic		Mild Dyslexic		Moderate Dyslexic	
	M	***SD***	***M***	***SD***	***M***	***SD***
LS of 1.5, LL of 60 – 70cpl	208.58	35.09	198.06	32.09	232.80	29.63
LS of 1.5, LL of 80 – 90cpl	206.10	37.72	192.81	22.72	233.64	25.24
LS of 2.0, LL of 60 – 70cpl	209.93	36.31	192.49	21.46	231.09	21.54
LS of 2.0, LL of 80 – 90cpl	202.33	31.52	191.84	28.95	224.50	21.61

Note. 'LS' denotes Line Spacing; 'LL' denotes Line Length. Measures in milliseconds (ms).

Table 4.6 Three-way mixed ANOVA with repeated measures on average fixation duration

Effects		***F***	***p***	η_p^2
Main	Group	$F(2, 21) = 3.69$	.04	.26
	Line Spacing	$F(1, 21) = 1.94$	ns	.09
	Line Length	$F(1, 21) = 1.95$	ns	.09
Two-way interaction	Group × Line Spacing	$F(2, 21) = 0.26$	ns	.02
	Group × Line Length	$F(2, 21) = 0.08$	ns	.01
	Line Spacing × Line Length	$F(1, 21) = 0.37$	ns	.02
Three-way interaction	Group × Line Spacing × Line Length	$F(2, 21) = 0.72$	ns	.06

Note. 'ns' indicates not statistically significant.

There was a significant main effect for Group with large effect size, $F(2, 21) = 3.69$, $p = .04$, $\eta_p^2 = .26$. This is illustrated in Figure 4.7. A post-hoc analysis showed that moderate dyslexic participants ($M = 230.51$, $SD = 23.78$) had significantly longer fixation duration than mild dyslexic participants ($M = 193.80$, $SD = 25.47$), while the other differences were not significant.

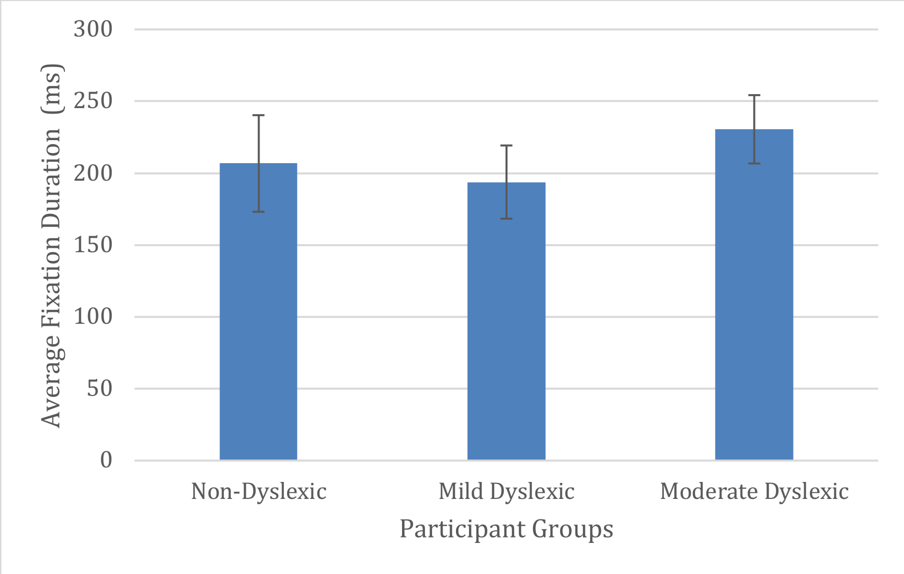


Figure 4.7 Mean average fixation duration for the three participant groups

4.3.3.2 Reading Performance

a. Text Reading Time

Table 4.7 shows the descriptive statistics for the text reading time for the three participant groups. Text reading time were normally distributed as assessed by Shapiro-Wilk's test of normality. The assumption of homogeneity of variances was met except Line Spacing 1.5 with 60 – 70cpl (F = 4.03, p = .03) and Line Spacing 2.0 with 80 – 90cpl (F = 6.14, p = .01), as assessed by Levene's test for equality of variances. Mauchly's test of sphericity indicated that the assumption of sphericity was met. The results of the ANOVA are summarised in Table 4.8.

Table 4.7 Descriptive statistics for reading time (ms) for all participant groups

	Non-Dyslexic		**Mild Dyslexic**		**Moderate Dyslexic**	
	M	***SD***	***M***	***SD***	***M***	***SD***
LS of 1.5, LL of 60 – 70cpl	70220.73	16577.29	71952.10	24907.26	111095.29	39319.31
LS of 1.5, LL of 80 – 90cpl	65836.00	16579.33	68018.58	16515.97	94863.00	20868.13
LS of 2.0, LL of 60 – 70cpl	65198.21	15129.57	73036.70	20160.59	92903.63	24990.59
LS of 2.0, LL of 80 – 90cpl	64389.33	10139.43	68099.83	15187.37	96083.90	27051.89

Note. 'LS' denotes Line Spacing; 'LL' denotes Line Length. Measures in milliseconds (ms).

Table 4.8 Summary of three-way mixed ANOVA on reading time

Effects		*F*	*p*	η_p^2
Main	Group	$F(2, 21) = 6.12$	.01	.37
	Line Spacing	$F(1, 21) = 3.03$	ns	.13
	Line Length	$F(1, 21) = 5.00$	.04	.20
Two-way interaction	Group × Line Spacing	$F(2, 21) = 1.52$	ns	.13
	Group × Line Length	$F(2, 21) = 0.32$	ns	.03
	Line Spacing × Line Length	$F(1, 21) = 4.01$	ns	.16
Three-way interaction	Group × Line Spacing × Line Length	$F(2, 21) = 2.85$	ns	.21

Note. 'ns' indicates not statistically significant.

There was a significant main effect for Group with large effect size, $F(2, 21) = 6.12$, $p = .01$, $\eta_p^2 = .37$. Figure 4.8 illustrates this interaction. Post-hoc analysis indicated that moderate dyslexic participants ($M = 98736.46$, $SD = 28418.06$) had significantly longer reading time than mild dyslexic participants ($M = 70276.80$, $SD = 18728.86$), $p = .03$ and non-dyslexic participants ($M = 66411.07$, $SD = 14292.98$), $p = .01$.

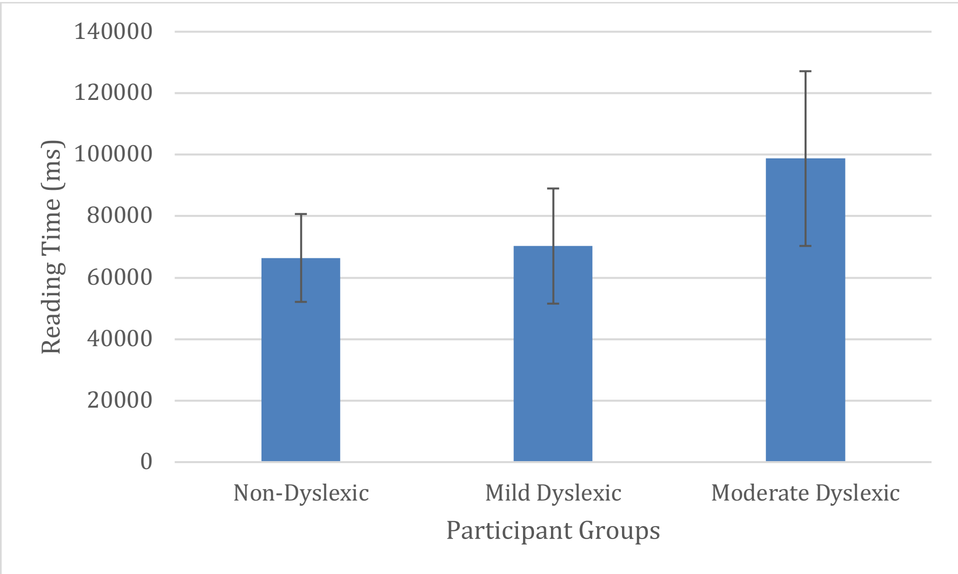


Figure 4.8 Mean text reading time for the three participant groups

There was also a significant main effect for Line Length with large effect size, $F(1, 21) = 5.04$, $p = .02$, $\eta_p^2 = .20$, where line length of 60 – 70cpl ($M = 80734.44$, $SD = 28573.47$) had significantly longer reading time than 80 – 90cpl ($M = 76215.11$, $SD = 22276.71$).

b. Comprehension Scores

Participants answered a total of 12 comprehension questions across the 4 texts. The mean of total comprehension score for non-dyslexic participants (M = 7.88, SD = 1.81) was lower than mild dyslexic participants (M = 10.5, SD = 1.07) and moderate dyslexic participants (M = 10.1, SD = 1.73). The total comprehension scores were normally distributed as assessed by Shapiro-Wilk's test for normality. There were no violations on homogeneity of variances assessed by Levene's test for equality of variances.

There was a significant effect for Comprehension Score with large effect size, $F(2, 21)$ = 6.54, $p < .01$, $\eta_p^2 = .38$. This is illustrated in Figure 4.9. A post-hoc analysis showed that non-dyslexic participants had significantly lower comprehension scores than mild dyslexic participants ($p < .001$) and moderate dyslexic participants ($p = .03$).

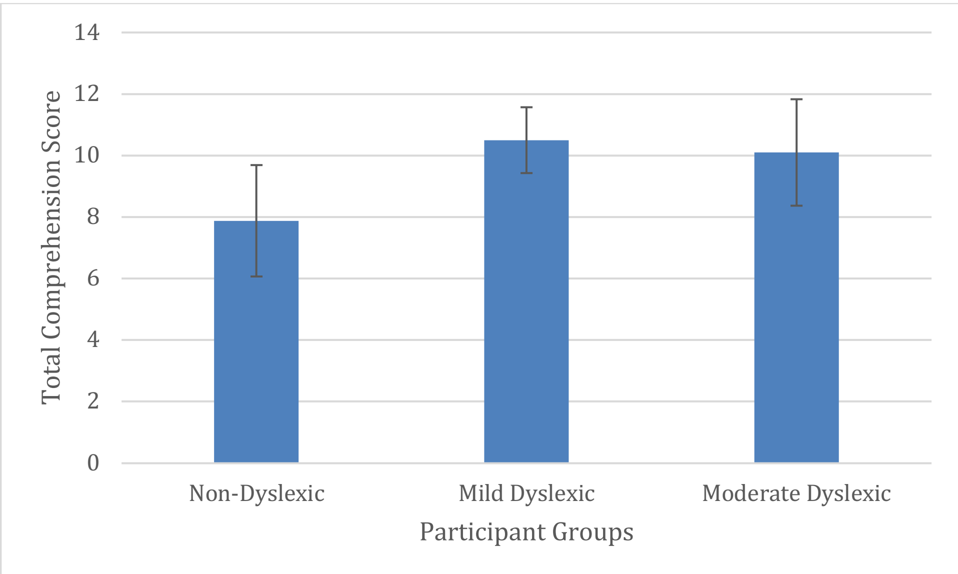


Figure 4.9 Mean total comprehension score for the three participant groups

To investigate the relationship between total comprehension score and two line spacings and two line lengths, Chi-square tests were conducted. There was no statistically association between comprehension scores with Line Spacing ($X^2(1) = 0.10$ $p = .76$) and Line Length ($X^2(1) = 2.23$, $p = .13$).

4.3.3.3 Participants' Preferences

a. Ease of Reading

Table 4.9 shows the descriptive statistics of ratings given for ease of reading for all combinations of line spacing and line length. It can be seen line spacing of 2.0 had the highest rating of ease of reading with 60 – 70cpl (*M* = 5.29, *SD* = 1.71) and 80 – 90cpl (*M* = 5.29, *SD* = 1.23). Note that lower ratings of ease of reading indicate greater difficult of reading while higher ratings indicate greater ease of reading. Figure 4.10 visualises the means and standard deviations of the ratings.

Table 4.9 Ease of reading ratings for combinations of line spacing and line length

Line Spacing	Line Length	Group	*M*	*SD*
1.5	60 - 70cpl	Non-Dyslexic	4.62	1.19
		Mild Dyslexic	4.87	1.55
		Moderate Dyslexic	3.62	1.41
		Total	**4.37**	**1.44**
	80 - 90cpl	Non-Dyslexic	4.25	1.04
		Mild Dyslexic	4.87	0.64
		Moderate Dyslexic	4.87	1.46
		Total	**4.67**	**1.09**
2.0	60 - 70cpl	Non-Dyslexic	5.12	1.81
		Mild Dyslexic	5.50	1.51
		Moderate Dyslexic	5.25	1.98
		Total	**5.29**	**1.71**
	80 - 90cpl	Non-Dyslexic	5.87	0.64
		Mild Dyslexic	5.37	0.74
		Moderate Dyslexic	4.62	1.77
		Total	**5.29**	**1.23**

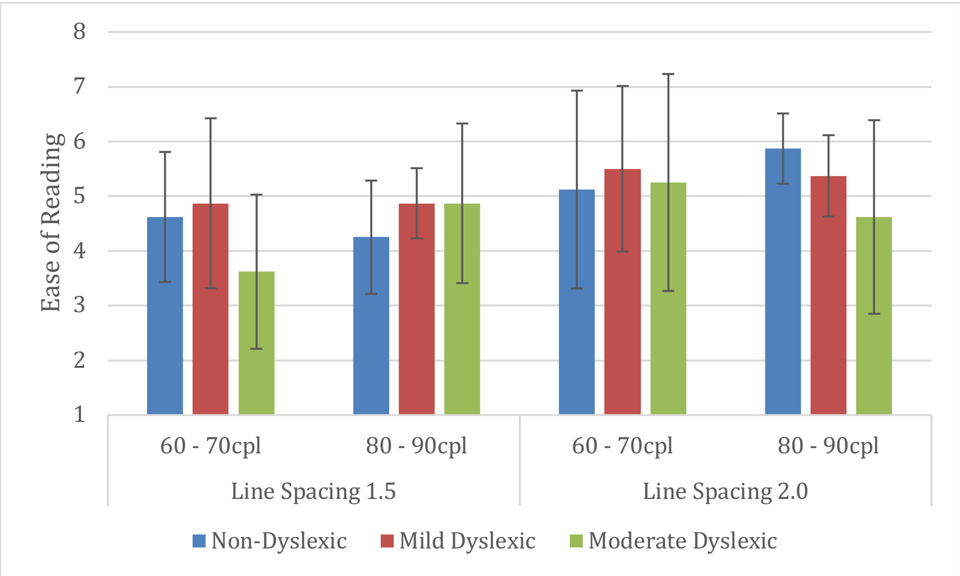


Figure 4.10 Means and standard deviations for ease of reading ratings for different line spacing and line length for the three participant groups

Ratings for ease of reading were normally distributed except for three groups: Non-dyslexic with Line Spacing 2.0 and 80 – 90cpl, $p = .04$; Mild dyslexic with Line Spacing 2.0 and 80 – 90cpl, $p = .03$; and Mild dyslexic with Line Spacing 1.5 and 80 – 90cpl, $p = .04$, as assessed by Shapiro-Wilk's test of normality. The assumption of homogeneity of variances was met except Line Spacing 2.0 with 80 – 90cpl ($F = 7.56$, $p < .01$), as assessed by Levene's test for equality of variances. Mauchly's test of sphericity indicated that the assumption of sphericity was met. The results of the ANOVA are summarised in Table 4.10.

Table 4.10 Summary of three-way mixed ANOVA on ease of reading

Effects		***F***	***p***	η_p^2
Main	Group	$F(2, 21) = 1.74$	ns	.14
	Line Spacing	$F(1, 21) = 4.58$	.04	.18
	Line Length	$F(1, 21) = 0.20$	ns	.01
Two-way interaction	Group × Line Spacing	$F(2, 21) = 0.17$	ns	.02
	Group × Line Length	$F(2, 21) = 0.12$	ns	.01
	Line Spacing × Line Length	$F(1, 21) = 0.96$	ns	.04
Three-way interaction	Group × Line Spacing × Line Length	$F(2, 21) = 8.55$	<.01	.45

Note. 'ns' indicates not statistically significant.

There was a significant main effect was for Line Spacing with large effect size, $F(1, 21) = 4.58$, $p = .04$, $\eta_p^2 = .18$, where line spacing of 2.0 ($M = 5.29$, $SD = 1.47$) had significantly higher ease of reading than line spacing of 1.5 ($M = 4.52$, $SD = 1.27$).

There was a significant three-way interaction for Group, Line Spacing and Line Length with large effect size, $F(2, 21) = 8.55$, $p < .01$, $\eta_p^2 = .45$. Figure 4.10 illustrates this interaction. Further post-hoc analysis indicated that this three-way interaction was due to the presence of significant simple tw0-way interaction, Line Spacing and Line Length for non-dyslexic participants ($F(1, 7) = 7.99$, $p = .03$, $\eta_p^2 = .53$) and moderate dyslexic participants ($F(1, 7) = 15.29$, $p < .01$, $\eta_p^2 = .69$), but not for mild dyslexic participants ($F(1, 7) = 0.04$, $p = .85$, $\eta_p^2 = .01$).

Further post-hoc analysis on non-dyslexic participants indicated that this simple two-way interaction was due to the presence of significant simple main effect of Line Spacing when reading with 80 – 90cpl, $F(1, 7) = 9.32$, $p = .02$, $\eta_p^2 = .57$, but not with 60 – 70cpl. Further post-hoc analysis was made on non-dyslexic participants reading in 80 – 90cpl condition where all pairwise comparisons were run between different line spacing trials.

The mean rating for ease of reading with line spacing of 1.5 (M = 4.25, SD = 1.04) was significantly lower than line spacing of 2.0 (M = 5.87, SD = 0.64).

Further post-hoc analysis on moderate dyslexic participants indicated that the simple two-way interaction had no statistically significant simple main effect for either Line Spacing or Line Length. It can be seen from Figure 4.10, moderate dyslexic participants reading in 60 – 70cpl had lowest mean rating for ease of reading with line spacing of 1.5 (M = 3.62, SD = 1.41) than line spacing of 2.0 (M = 5.25, SD = 1.98). However, the standard deviations were large.

b. Preferred Line Spacing and Line Length

Table 4.11 - 4.12 shows the frequency of preferred line spacing and line length for each participant group. To investigate participants' preferred combination of line spacing and line length, Chi-square independence tests were conducted between the three participant groups with line spacing and line length of their stated preferred combinations. There was no statistically association between participant groups and the line spacing of their preferred combination, $X^2(2, n = 24) = 0.38, p = .82$; and line length of their preferred combination, $X^2(2, n = 24) = 2.34, p = .31$.

Table 4.11 Distribution of preferences for line spacing by the three participant groups

Participant Group	N	Line Spacing	
		1.5	2.0
Non-Dyslexic	8	2	6
Mild Dyslexic	8	3	5
Moderate Dyslexic	8	3	5
Total	24	8	16

Table 4.12 Distribution of preferences for line length by the three participant groups

Participant Group	N	Line Length	
		60 – 70cpl	80 – 90cpl
Non-Dyslexic	8	3	5
Mild Dyslexic	8	6	2
Moderate Dyslexic	8	4	4
Total	24	13	11

To investigate the distribution of line spacing or line length preferred by all participants together, Chi-square likelihood tests were conducted. For Line Spacing, the distribution of preferences was not different from a random distribution, $X^2(1, n = 24) = 2.67, p = .10$.

For Line Length, the distribution of preferences was not different from a random distribution, $X^2(1, n = 24) = 0.17, p = .68$.

4.3.4 Discussion and Conclusion

The study investigated the effects of line spacing and line length on eye gaze behaviour, reading performance and preference measures on detailed text reading from a computer screen using eye tracking equipment. This study compared adult English native speakers with and without dyslexia. Participants were asked to complete an online dyslexia checklist (the ARQ) to screen them into non-dyslexic, mild and moderate dyslexic. This study measured two levels of line spacing: 1.5 and 2.0; and two levels of line length: 60 – 70cpl and 80 – 90cpl.

Overall, this study found that line spacing had significant effect on rating of ease of reading and no significant effect on eye gaze behaviour and reading performance measures. Line length had significant effects on number of fixations and reading time. The three participant groups had significant differences in fixation duration and reading time. Details for each of these aspects are discussed below.

4.3.4.1 Effects of Line Spacing

According to various guidelines for text presentation for people with dyslexia, line spacing between 1.5 and 2.0 are recommended (British Dyslexia Association, 2018b; de Santana et al., 2012; Evett & Brown, 2005), although Rello, Kanvinde, et al. (2012b) recommended only 1.4. It is important to note that most of these guidelines were not based on empirical evidence. However in recent guidelines based on empirical work, Rello & Baeza-Yates (2015) have proposed using the default line spacing of 1.0 since they found no effect on fixation duration and preferences for both dyslexic and non-dyslexic participants.

In a study of quantifying vertical critical spacing on non-disabled users, no additional benefit was found when using larger than line spacing of 1.5 (Chung, 2004). Critical spacing is a distance where flanker letters (letters that distract reading target letter) starts to degrade reading performance. This present study found that both line spacings of 1.5 and 2.0 had no effect on fixation duration, and also had no effect on other measurements such as number of fixations, reading time and comprehension scores. Based on Chung's (2004) finding, different significant result might be yielded if the present study tested line spacing smaller than 1.5.

Regarding participants' preferences, both line spacing 1.5 and 2.0 were equally preferred. However, line spacing 2.0 was rated as easier to read than 1.5. Other researchers (Rello, Pielot, et al., 2013; Schoonewelle, 2013) have found similar preference results in which bigger line spacing was preferred, although the spacing did not have significant effects on comprehension score (Rello, Pielot, et al., 2013; Schoonewelle, 2013) and average fixation duration (Rello, Pielot, et al., 2013).

4.3.4.2 Effects of Line Length

Early work with printed materials on line length for non-disabled readers led to the recommendation of 70cpl (Dyson & Haselgrove, 2001). For online materials, 65 – 75cpl was recommended (Bernard, Fernandez, Hull, & Chaparro, 2003). Guidelines for people with dyslexia recommended line length between 60 to 70cpl (British Dyslexia Association, 2018b; de Santana et al., 2012; Evett & Brown, 2005), albeit no empirical evidence was provided to support this recommendation. Furthermore, there is a lack of information about whether this recommended line length is to be use for print or computer presentations. To date, limited studies have provided empirical data on line length on computer screens for participants with dyslexia (Rello & Baeza-Yates, 2015; Rello et al., 2012b). Rello, Kanvinde, et al. (2012b) recommended 77cpl, however in her recent guidelines, no specific recommendation was made (Rello, 2014b; Rello & Baeza-Yates, 2015).

Surprisingly, the present study found the length of 60 – 70cpl had significantly higher number of fixations and took longer time to read for all participant groups. In addition, there was no correlation between comprehension scores and line length. On the opposite, participants read better using longer line length (80 – 90cpl) which they were faster, had fewer fixations and no effect on comprehension. Other researchers have found similar results in which longer line lengths were faster to read (Rello et al., 2012b) at no cost on comprehension (Duchnicky & Kolers, 1983; Dyson & Haselgrove, 2001; Dyson & Kipping, 1998; Shaikh, 2005). However with a different device (iPod), Schneps et al. (2013) found that dyslexic participants had shorter reading times, a fewer number of fixations and no effect on comprehension when reading a very short line length (12.7cpl) compared to a longer line length (67.2cpl). As all participants in this present study were university students and familiar with reading from web pages, thus they are used to the long line lengths typically found on these pages. Studies with college students have shown that they have faster reading speed and are efficient at reading with longer line lengths (e.g. Shaikh, 2005).

Rello, Kanvinde, et al. (2012b) found line length of 88cpl (the longest length used in that study) had the shortest fixation durations, yet no statistical test. However, her recent guidelines shown that line length was not statistically significant on fixation duration and preferences (Rello & Baeza-Yates, 2015). Similar results were found in the present study which both line length (60 – 70cpl and 80 – 90cpl) had no significant effects on fixation duration and preferences.

4.3.4.3 Overall Effects of Participant Group

Regarding the differences between the participant groups, no other study compared participants without dyslexia, with participants with both mild and moderate dyslexia separately. Therefore, no comparison between these three groups could be made with other studies. There is good evidence that dyslexics show poorer eye gaze behaviour in reading, with more fixations and longer fixation duration (Rayner, 1998). In line with these findings, this present study found that moderate dyslexic participants had a significant longer fixation duration but not more fixations and longer reading times. This shows that severity of dyslexia had some effect on eye gaze and reading performance in their participants.

For comprehension score, people with dyslexia can read relatively well and understand what they have read even though it is generally accepted that they may have problems in comprehension (Snowling, 1998, 2012; Vellutino et al., 2004). For non-dyslexic participants, there was a trade-off between their comprehension and time, where they had lower score but shorter fixation duration and reading time than mild and moderate dyslexic participants. As found in other studies, when reader read faster, they tend to have lower comprehension scores (Dyson & Haselgrove, 2000, 2001).

4.3.4.4 Limitations of this Study

As discussed in Section 3.3.4.4, most of the limitations in previous study were overcome in this present study. Participants in this present study read a smaller number of texts and also number of words in the texts were shorter in order to reduce the overall duration of the experiment. The theme of the texts was change from Olympics (in Study 1) to variety of topics (fruit, vegetable and cities) that might gather more interest from participants. The texts were counterbalanced across all line spacing-line length conditions to avoid the results being influenced by language variables. In addition, lights and windows in the room used in this study were controlled in order to minimise the number of participants with erratic eye gaze results. In the case of erratic eye gaze behaviour, the participant was fully replaced by another one.

Another limitation of this study is most of the participants were university students. As university students, they may have developed strategies in reading texts with a variety of line lengths and line spacings. Therefore, this means that caution must be exercised when generalising findings from this present study.

4.3.4.5 Conclusions

This study has shown that moderate dyslexic participants had longer fixation and longer reading time than non-dyslexic and mild dyslexic participants. It also showed that line spacing had an effect on subjective measures (participant preferences) and line length had an effect on objective measures (eye gaze behaviour). For line spacing, participants felt that larger spacing of 2.0 was easier to read. On the other hand, they felt that line length had no effect on ease of reading. However from their eye gaze data (objective measures), line length showed significant differences whereby all participants, including participants with mild and moderate dyslexia, had better reading behaviour with longer line length of 80-90cpl. Even though this study had a small sample size, the effect sizes of the significant results were large. Therefore, the significant results were meaningful. Overall, it can be seen that objective results support longer line length are better for all participants while subjective results support bigger line spacing are easier to read for all participants.

CHAPTER 5: STUDY 3 – EFFECTS OF MENU ORGANISATION AND VISIBILITY ON WEB NAVIGATION

5.1 Introduction

This chapter describes Study 3 in this programme of research which aims to answer Research Questions 2(a) and 2(b), discussed in Chapter 1:

2 (a) To what extent does the organisation of the web navigation used in web navigation affect eye gaze behaviour, navigation performance and opinions of adults with dyslexia compared to adults without dyslexia?

2 (b) To what extent does the visibility of the web navigation used in web navigation affect eye gaze behaviour, navigation performance and opinions of adults with dyslexia compared to adults without dyslexia?

As discussed in Chapter 1, there are varying degree of severity for people with dyslexia. Studies in the text presentation (discussed in Chapter 3 and Chapter 4) categorised participants of dyslexia based on their severity, mild and moderate dyslexia. However, no major significant differences between them were found. Therefore, this study will only include non-dyslexic and dyslexic participant groups.

As was seen from the discussion in Chapter 2, there is very limited research on the problems that people with dyslexia might encounter in navigating websites (Al-Wabil et al., 2007, 2008). In addition, insights discovered from these studies were gathered from a very limited number of participants. Scarcity of research showed that a lack of attention has been given to this area. According to Al-Wabil et al. (2007), highly textual content and poor navigation structure can be difficult for people with dyslexia. On the World Wide Web, menus for navigation are an important element that help Web users to locate and retrieve information. Both having difficulty in reading, and an inability to locate information in websites can frustrate people with dyslexia. Therefore, designing navigation menus appropriately is imperative to help reduce the effort of people with dyslexia. It is worth noting that the navigation studies with people with dyslexia discussed in Chapter 2 had different purpose from this study which Al-Wabil et al. (2007) explored range of difficulties in navigating content in a web site while Al-Wabil et al. (2008) focused on visual attention when given different types of tasks.

In this study, navigation menus were arranged with different organisations and levels of visibility. In order to minimise participants' frustration in finding information, tasks were therefore assessed in a preparatory study. This chapter will present the method and results of the preparatory study. Materials finalised from the preparatory study will then be used in the Study 3. The rest of this chapter will also describe the method, results and discussion of Study 3.

5.2 Preparatory Study – Assessment of Question Difficulty

According to Rosenfield and Morville (as cited in Leuthold, Schmutz, Bargas-Avila, Tuch, & Opwis, 2011), users are able to handle large numbers of menu items if the menus are easy to scan. In addition, the working memory of people with typical cognitive abilities can store around 7 ± 2 items at a time (Miller, 1956). By considering all of the above, a web site was developed to be used in the preparatory study which will have less than nine items in its sub-menus. The web site will be used later in Study 3, which involves different organisations and visibilities of menus including dynamic-sub-menus.

Nine multiple choice questions were developed to be used with the experimental web site. The preparatory study aimed to measure the difficulty of the questions using the menus and participants' familiarity on the questions.

Initially, participants were asked to read the proposed questions and navigate the experimental web site to find the answer for each of the questions. In addition to answering the question, they were asked to rate the difficulty of finding the answer. This enabled me to make a selection of which questions that would be used in Study 3.

5.2.1 Method

5.2.1.1 Design

A repeated measures design was chosen to investigate the difficulty of finding the answer to the questions (the Answer Difficulty Rating), how accurately participants could answer the questions (the Correct Answer) and whether they already knew the answer for the questions before navigating the experimental web site (the Previous Knowledge Assessment).

Accuracy was measured by participants' correct answers to the multiple choice question. Participants' familiarity with the subject matter of the questions as measured by asking whether they already knew the answer before navigating the web site.

5.2.1.2 Participants

Participants were all English native speakers. Five participants took part, including one female participant. Their ages ranged from 27 to 66 years old.

5.2.1.3 Materials

An experimental web site about tourism in Canada and questions were developed which would be used for the experimental tasks in Study 3. Each question had three response alternatives. The web site was adapted from Lonely Planet Canada web site[20] and Lonely Planet Canada Guide (2017). Content from Lonely Planet was chosen as it is a reputable source for tourism and had higher customer ratings in Amazon.co.uk (4.9/5.0 stars), Amazon.com (3.7/5.0 stars) and goodread.com (4.0/5.0 stars). Menus were organised such that menu bars are located on top of the web page and menu items were visible all the time. There were nine items in the main menus and the maximum number of six items in the sub-menus. Figure 5.1 shows a web page in the experimental web site used in the preparatory study. Figure 5.2 shows the hierarchical structure of the menus used the web site. The hierarchical structure was adapted from Lonely Planet (2017).

As discussed in Section 3.2, questions were developed based on work by Davey (1988), using only straight-forward questions with the answer explicitly stated in the passage. Each question had one correct answer and two alternative distractor choices. Figure 5.3 shows one of the multiple choice questions and answers. Answers for each of the questions can be found in different web pages.

Each of the questions was accompanied by a difficulty rating and a familiarity question. The Answer Difficulty Rating was rated after participants found the answer for the question in the web site. The ratings were made on 9-point Likert items. Participants' familiarity with the subject matter of the questions was measured by asking whether they already knew the answer before navigating the web site using a simple Yes/No

[20] https://www.lonelyplanet.com/canada

option. Appendix M shows the nine questions with its accompanying difficulty rating and familiarity question.

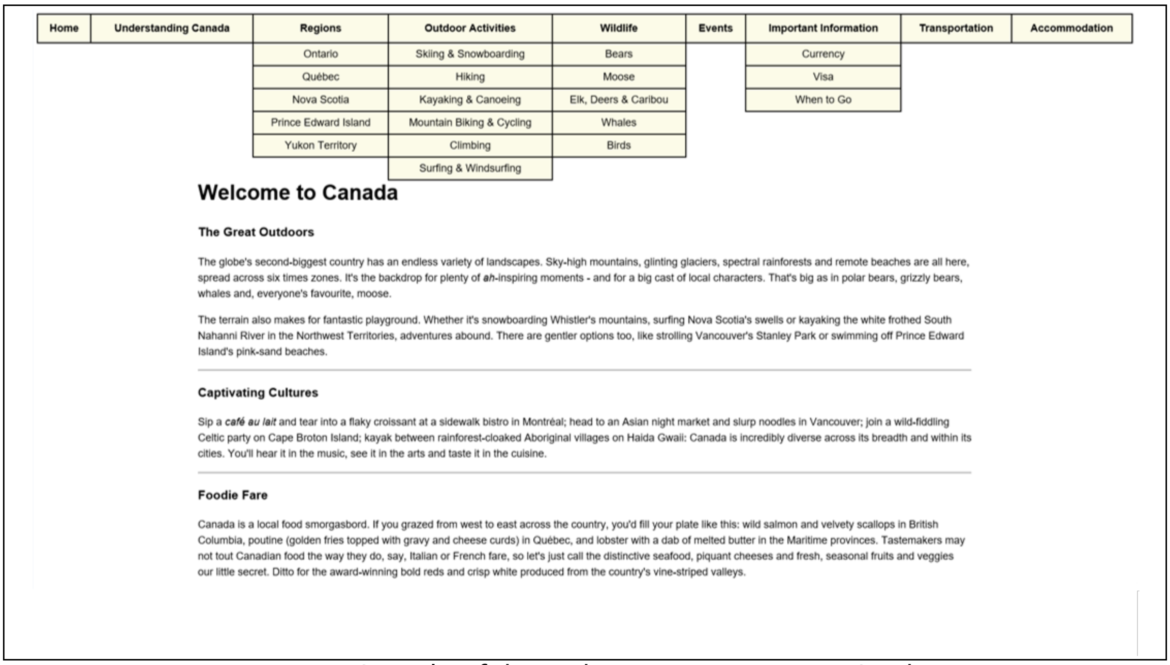


Figure 5.1 Sample of the web site in Preparatory Study 3

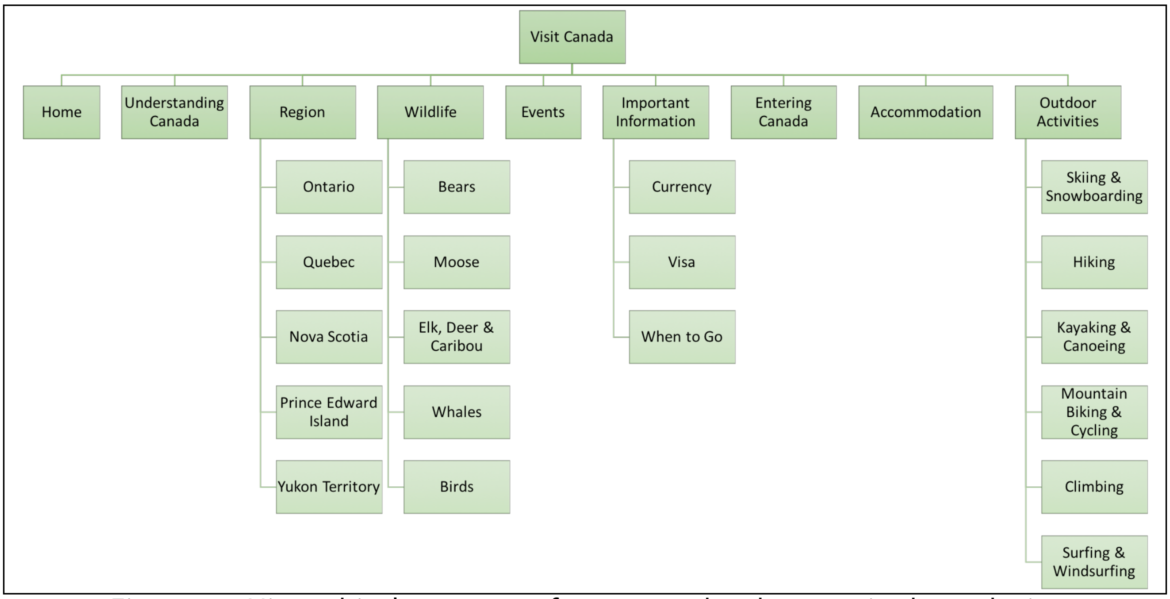


Figure 5.2 Hierarchical structure of menus and web pages in the web site

How much does an Australian visitor need to pay for an electronic Travel Authorization (eTA)?

a. $7
b. $100
c. US$6

Figure 5.3 Sample of multiple choice questions with its alternatives

5.2.1.4 Procedure

An email was sent to participants containing the link to the web site and a document file containing all the multiple choice questions and other questions to be answered. They were asked to read the question, open the link to the web site and find the answer by navigating through the web site. They then rated the difficulty of process in finding the answer for that question on a 9-point Likert scale item. They also needed to indicate if they already knew the answer before navigating to the web site, using the Yes/No option. After they had answered the questions and difficulty assessments, participants returned the file through email or on paper.

5.2.2 Results and Discussion

For each question, participants' accuracy scores were averaged into Correct Answer, participants' difficulty ratings in finding the correct answer in the web site were averaged into Answer Difficulty Rating, and participants' familiarity on the subject matter were averaged into Previous Knowledge Assessment.

A total of four questions was needed for Study 3, the correct answer for two of the questions can be answered from pages accessed via the main menu in the web site and the other two can be answered from pages accessed via sub-menus. A number of factors contributed to the choice of questions. The question was discarded if it was too difficult, too easy or most of participants had previous knowledge about the topic. Questions were aimed to have an Answer Difficulty Rating within the range of 5.00 to 8.75 and a Previous Knowledge Assessment within the range of 0.00 to 0.50. Since all of the questions were within the criteria range, I decided to have a balanced number of difficult (lower rating for Answer Difficulty Rating) and easy questions (higher rating for Answer Difficulty Rating).

Question 2 and Question 4 were selected to represent easy questions, while Question 8 and Question 9 were selected to represent difficult questions. Correct answers for Question 2 and Question 8 can be answered from pages accessed via main menu, while correct answers Question 4 and Question 9 can be answered from pages accessed via sub-menus.

Table 5.1 shows the results for all questions in the Preparatory Study. From the table, location of the correct target (whether participants should click on main menu or sub-menu) was shown; Correct answer was either 0 (incorrect answer) or 1 (correct answer);

Answer Difficulty Rating ranged from 0 (very difficult) to 8 (not difficult at all); and Previous Knowledge was either 0 (no) or 1 (yes). Question marked in red indicates it was discarded from being used in Study 3. Means and standard deviations for all questions that have been selected was also shown in the table.

Table 5.1 Result for all questions in the Preparatory Study

Question	**Location of the Correct Target**	**Correct Answer (*M*)**	**Answer Difficulty Rating (*M*)**	**Previous Knowledge Assessment (*M*)**
1	Sub-Menu	1.00	7.00	0.00
2	Main Menu	1.00	7.60	0.00
3	Main Menu	1.00	6.20	0.00
4	Sub-Menu	1.00	7.00	0.00
5	Sub-Menu	0.40	8.20	0.00
6	Main Menu	1.00	8.20	0.00
7	Sub-Menu	0.60	7.40	0.00
8	Main Menu	1.00	5.80	0.00
9	Sub-Menu	1.00	5.80	0.20
Selected questions, *M (SD)*		1.00 (0.00)	6.55 (0.90)	0.05 (0.10)

Note. Questions marked in red were discarded.

From the selected questions in Table 5.1, all participants answered all questions accurately ($M = 1.0$). Participants had the most difficulty in find answers for Questions 8 and 9 ($M = 5.80$), while Question 2 ($M = 7.60$) had the easiest rating. One participant was familiar with Question 9 ($M = 0.20$) since he had travelled to Canada before participating in this preparatory study. However, none of the participants familiar with Questions 2, 4 and 8 ($M = 0.00$).

The preparatory study has allowed me to create sets of questions with a balanced number of easy and difficult questions, which would not be familiar to participants. These sets will now be used in the Study 3.

5.3 Study 3

As discussed in Chapter 2, limited studies were found in the literature about navigation especially on how to design menus. Designing menus well is crucial to allow people to find information more easily on the Web, particularly for people with dyslexia as they may have greater challenges in using web search (Morris, Fourney, Ali, & Vonessen, 2018). Extensive research has been done on web navigation, particularly with non-disabled users on how to position menus (Bernard, Hamblin, & Chaparro, 2003; Burrell & Sodan, 2006; Fang & Holsapple, 2007; Faulkner & Hayton, 2011; James Kalbach &

Bosenick, 2003; Leuthold et al., 2011; McCarthy, Sasse, & Riegelsberger, 2004; Murano & Lomas, 2015; Murano & Oenga, 2012; Patsula, Detenber, & Theng, 2010; Pittsley & Memmott, 2012; Puerta Melguizo, Vidya, & van Oostendorp, 2012; Yu & Roh, 2002). In examining menu design for people with disabilities, limited research was found for this aspect of navigation (Harrysson, Svensk, & Johansson, 2004; Sevilla, Herrera, Martínez, & Alcantud, 2007; Williams & Hennig, 2015a, 2015b). In addition, very limited research was found on web navigation for people with dyslexia (Al-Wabil et al., 2007, 2008). This shows that research for people with dyslexia has overlooked this aspect of web design. This study will investigate whether different design of menus will influence the way people with dyslexia navigate a website, particularly the design of the organisation and the visibility of menus. For organisation, menus were organised into two different positions. Common organisations for menus are inverted-L shape menus (horizontally top of the page and vertically left of the page), however evidence from some studies has shown that users have better performance with menus placed horizontally across the top of the page (Murano & Oenga, 2012; Murano & Sander, 2016; Williams & Hennig, 2015b). For visibility, menus were organised into two different appearances. According to Nielsen (1999), dynamic menus are recommended. On the other hand, Fowler and Stanwick (2004) recommend displaying as many menu items as possible in logical groups at the same time on a single screen.

5.3.1 Method

5.3.1.1 Design

This study has a 2 × 2 × 2 between-participants design with participant group, organisation of menus and visibility of menus as the between participants variables. Participants were divided into two groups (non-dyslexic, dyslexic) on the basis of a diagnostic checklist for dyslexia (Snowling et al., 2012). As mentioned in earlier chapter, this study will only consider mild and moderate dyslexia as one group because no major differences were found between them. In addition, splitting them into different groups will require additional participants in this between-participants design. For the organisation variable, two levels were used: Unified and Fragmented. This variable related to how the menus were organised with either all menu items located on top of the page in Unified level (see Figure 5.5 and Figure 5.6), or the menu items organised on the top and left side of the page in the Fragmented level (see Figure 5.7 and Figure 5.8). For the visibility variable, two levels were used: Visible and Dynamic. This variable relates to the visibility of menu items, particularly for sub-menus with either all sub-menu items visible and static under each main menu in the Visible level (see Figure 5.5 and Figure 5.7), or the sub-menu items dynamically appearing when the participant hovered their mouse over in the Dynamic level (see Figure 5.6 and Figure 5.8).

The dependent variables were grouped into three measurements: (1) eye gaze behaviour, (2) navigation performance and (3) participants' opinions.

(1) The eye gaze behaviour measures are:
   a. entry time (time where participant starts looking at the area),
   b. dwell time (sum of duration from all fixations and saccades in the area)
   c. first fixation duration (duration of the first fixation that hit the area)
   d. number of revisits (total glances returned to the area)
   e. number of fixations (number of fixations made by participant in the area)
   f. average fixation duration (average of duration for one fixation made by participant in the area)

   The area mentioned above refers to the Area of Interest (AOI), meaning the location of particular elements of the webpage. All variables were repeated for three AOIs: target link, main menus and sub-menus. The number of fixations and number of revisits are good indicators of the importance of an element or its noticeability, dwell time reflects engagement of a participant with the area, while first fixation duration reflects the attention-getting property of an area (Bylinskii, Borkin, Kim, Pfister, & Oliva, 2015).

(2) The navigation performance measures are:
   a. navigation score (total of correct answers found for all questions)
   b. first mouse click times (time when participant first clicked on any menu item, whether it is the correct or wrong target)
   c. task completion times (time from when participant starts browsing the web site, to when they found the answer and closed the browser)

(3) The opinion measures are: 'ease of use', 'ease of remembering' and 'ease of learning'. All variables were gathered in a post-study questionnaire using a 7-point Likert scale items

Participants completed four experimental tasks in which they need to find the answer for a question about tourism in Canada, the answer to which was available on the experimental website. The order of tasks was counterbalanced.

Ethics approval for this study has been granted by Physical Sciences Ethics Committee of the University of York. Participants were ensured about the confidentiality and anonymity of their data.

5.3.1.2 Participants

Participants were asked to complete an online screening diagnostic tool (ARQ). They were categorised into non-dyslexic group if their diagnostic scores ranged between 0 to 10, and dyslexic group if their scores above than 10. A total of 64 participants took part in this study, 32 participants in each group (non-dyslexic and dyslexic). All participants were native English speakers with normal vision or vision correctable with eyewear.

The non-dyslexic group comprised 13 females and 19 males with ages ranging from 18 to 42 years (M = 22.25, SD = 4.59). Nineteen were undergraduate students (with two of them were working part time), five were Masters students (two of them were working part time), five were PhD students (one was working part time) and three were employed. All of them had more than 6 years of Web experience. Mean of Web expertise is 5.78 (SD = 1.07), on a 7-point Likert item.

The dyslexic group comprised 18 females and 14 males with ages ranging from 18 to 44 years (M = 20.61, SD = 4.70). Twenty-five were undergraduate students (three were working part time), six were Masters students (one was working part time) and one was employed. All of them had more than 6 years of Web experience. Mean of Web expertise is 5.32 (SD = 0.98), on a 7-point Likert item.

Participants were given a £10 gift voucher of Amazon to thank them for participating in the study.

5.3.1.3 Materials and Equipment

Equipment similar to that used in Studies 1 and 2 (laptop, computer screen, a remote eye tracker, wireless mouse and keyboard, cables and fixed chair) and dual monitor setup used in this study (see Section 3.3.1.3). The equipment and the setup for both participants and researcher were illustrated in Figure 3.4. Viewing distance between participants and the monitor screen was 60cm.

An online ARQ were used as a diagnostic tool to screen participants whether they were dyslexic or not (Snowling et al., 2012). The questionnaire (see Appendix C) asked participants information about their reading habits, literacy and any dyslexia symptoms. A pre-study questionnaire (see Appendix E) was used to collect information about participants' use of the web and their demographic information. The materials were discussed in detail in Section 3.3.1.3.

Two web sites (a practice web site and an experimental web site) and five navigational tasks (one practice task and four experimental tasks) were developed. Each task involved answering one multiple choice question. The practice web site is about insect and was adapted from Amateur Entomologists' Society[21] and ThoughtCo.[22,23,24,25]. There was only one version of the practice web site, in which the menus were organised on the left side of the web site with a dynamic sub-menus. Figure 5.4 illustrates the practice web site. The experimental web site used the same web site as in the preparatory study and had similar navigational structure, as discussed in Section 5.2 (see Figure 5.2). The experimental web site is about tourism in Canada and was adapted from Lonely Planet Canada web site[26] and Lonely Planet (2017). Four versions of the experimental web site were developed, one each for the levels of the organization variable and the visibility variable. Menus were either Unified (located on top of the page) or Fragmented (located on top and left side of the page). Sub-menus were either Visible (appeared all the time below particular item) or Dynamic (appeared when hovered with mouse). Figure 5.5 – Figure 5.8 illustrate web page from the different versions of experimental web sites with one menu items hovered with mouse (in orange colour).

---

[21] https://www.amentsoc.org/insects/
[22] https://www.thoughtco.com/how-honey-bees-communicate-1968098
[23] https://www.thoughtco.com/what-good-are-ants-1968090
[24] https://www.thoughtco.com/how-do-fireflies-light-1968122
[25] https://www.thoughtco.com/the-spider-life-cycle-1968557
[26] https://www.lonelyplanet.com/canada

The text passages in the web sites were presented left-justified, 80 – 90 characters per line (not including spaces), Arial typeface, 14-point size, and 1.5 line spacing. The passages in both practice and experimental web sites were presented as black text on a white background with limited images in some web pages. Menus in the practice web site had a very light grey (#F1F1F1, RGB (241, 241, 241)) with black text (#000000, RGB (0, 0, 0)). When hovering over with a mouse, the menu turned to very dark grey (#555555, RGB (85, 85, 85)) with white text (#FFFFFF, RGB (255, 255, 255)). Menus in the experimental web site have a linear gradient colour from light grey (#C1C1C1, RGB (193, 193, 193)) to very light grey (#F5F5F5, RGB (245, 245, 245)). When hovering over with a mouse, the menu item temporarily changed its background colour. The temporary colour used has a linear gradient colour from pure orange (#F8AC00, RGB (248, 172, 0)) to soft orange (#FAC754, RGB (250, 199, 84)). All menus, text passages and images fit in one screen, so no scrolling is required.

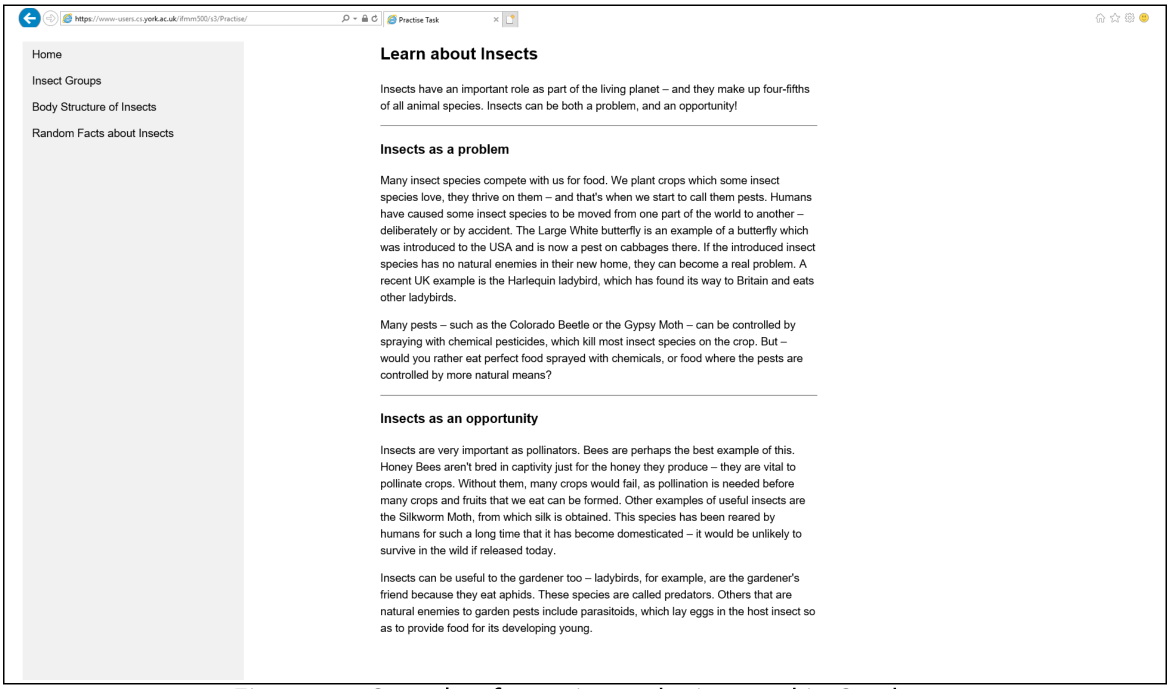


Figure 5.4 Sample of practice web site used in Study 3

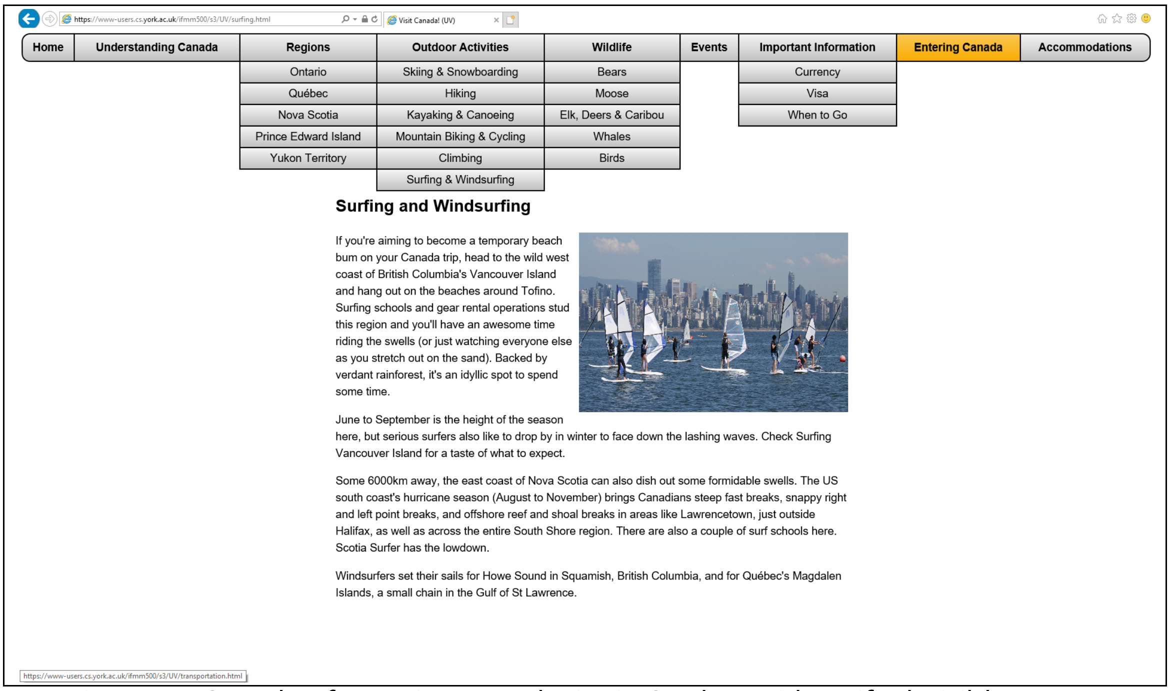


Figure 5.5 Sample of experiment web site in Study 3 with Unified-Visible menus

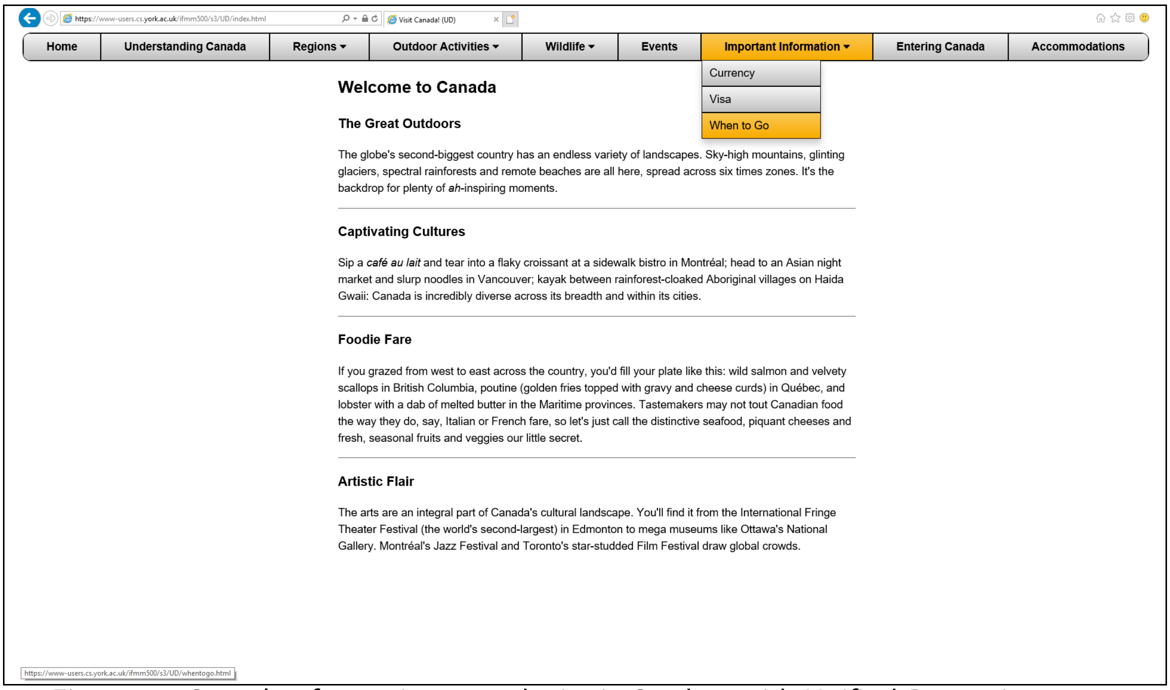


Figure 5.6 Sample of experiment web site in Study 3 with Unified-Dynamic menus

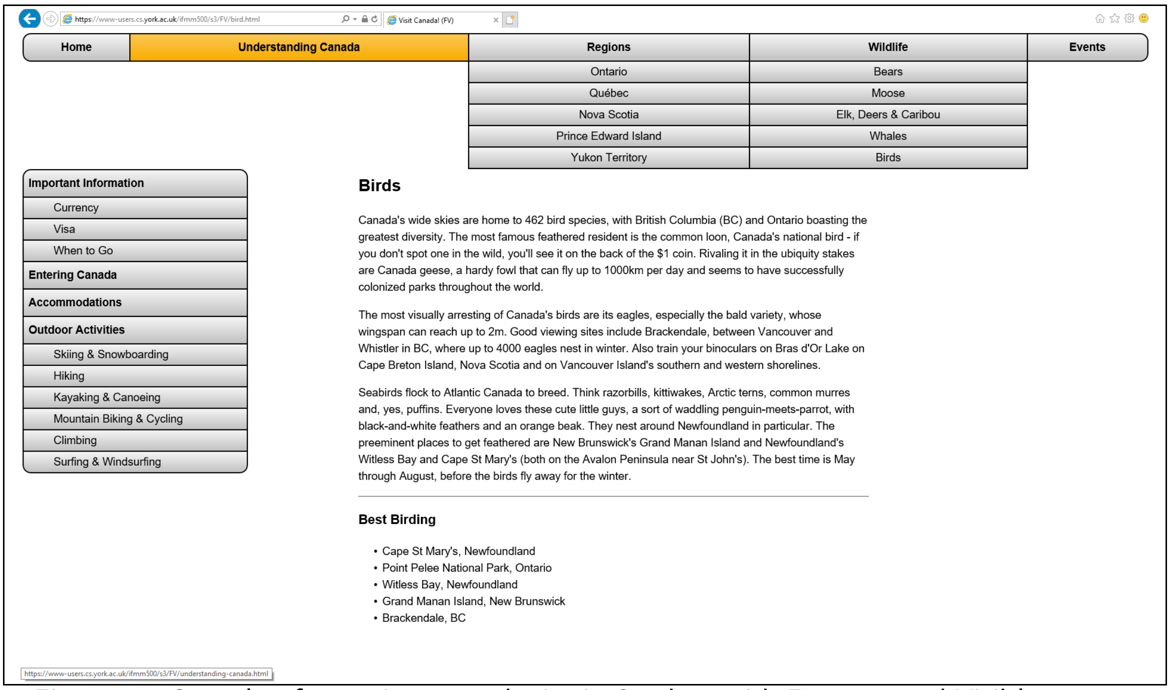


Figure 5.7 Sample of experiment web site in Study 3 with Fragmented-Visible menus

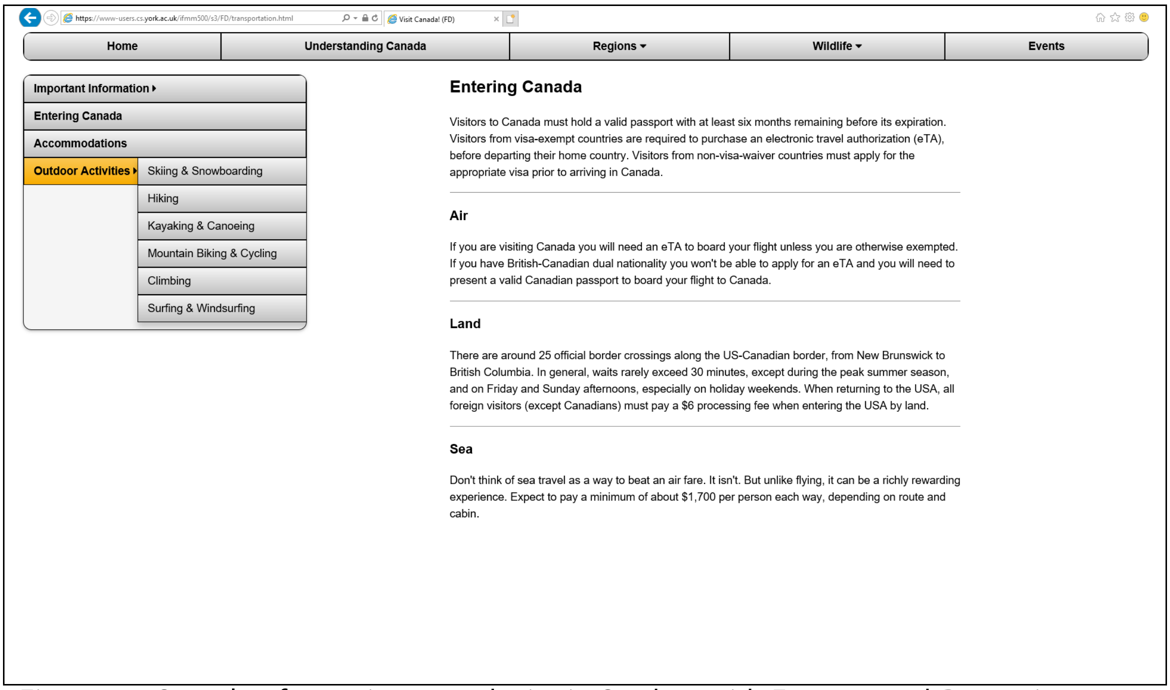


Figure 5.8 Sample of experiment web site in Study 3 with Fragmented-Dynamic menus

Multiple choice questions for the experimental web site were tested in the preparatory study, as discussed in Section 5.2. There was one question for the practice web site and four questions for the experimental web site. The question for practice web site was not included in the preparatory study, since it was only intended for practice purposes. Multiple choice questions were presented as black text on a white background while typeface, font size, line spacing and line length used default settings in the SMI

Experiment Center™. However, no information regarding the default setting was provided in the SMI Experiment Center™. Figure 5.9 shows a sample of a question for the experimental web site. Participants read the question at the beginning of the session, before they started navigating through the experimental web site to find an answer for the question. After navigating and close the web site, participants were presented the question again and now able to choose their answer. Figure 5.9 and Figure 5.10 shows the same question, in Figure 5.9 participants were able to read the question and in Figure 5.10 they were able to choose one possible correct answer. The 'Continue' button would be activated once participants made a selection of any of the options in radio button.

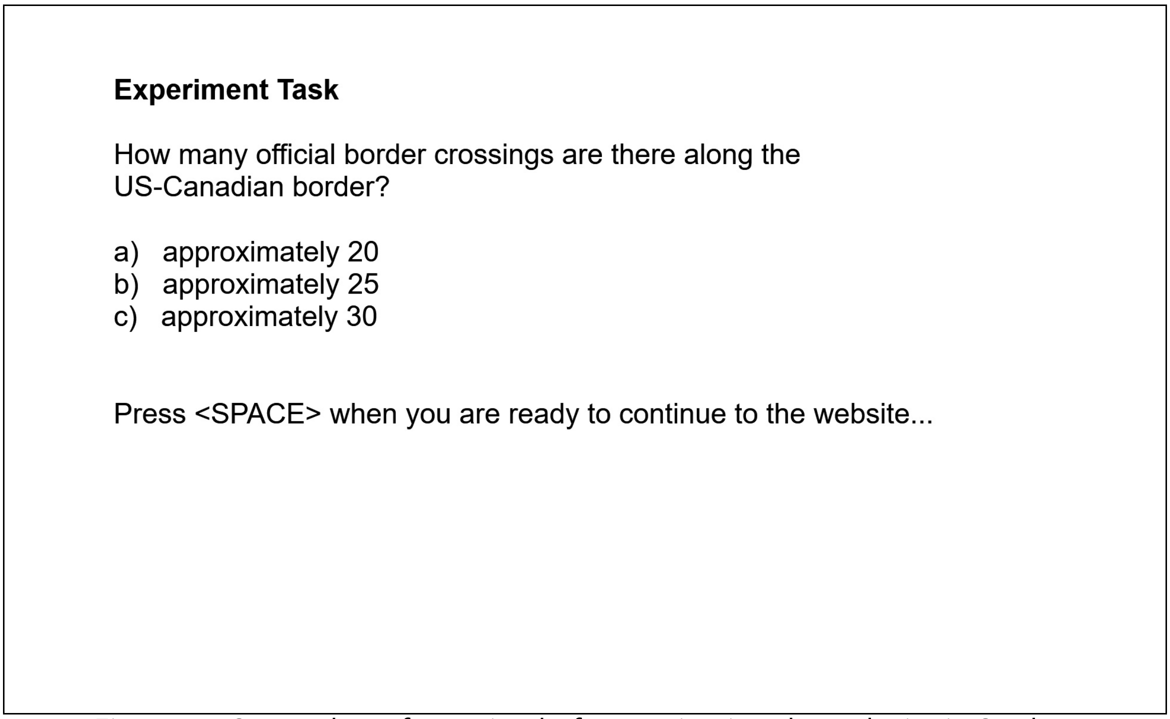


Figure 5.9 Screenshot of question before navigating the web site in Study 3

A small cue card (9cm × 12cm) was placed below the participants' monitor screen, to remind them of the question asked in the beginning of the task. They could look quickly at the cue card during the tasks. Cue cards were presented as black text on a white background, using Calibri 40-point size and line spacing of 1.0 in MS PowerPoint. Figure 5.11 shows the position of the cue card. There were five small cue cards used, on each for the practice task and the four experimental tasks. The cue card was changed in between the tasks.

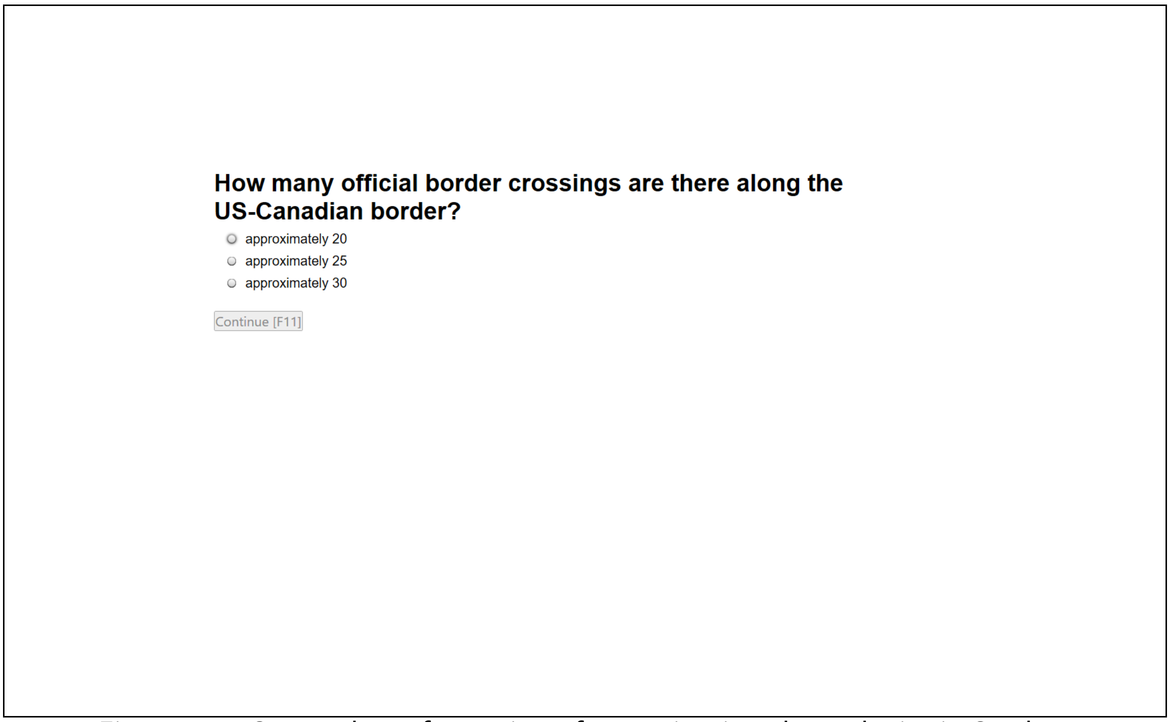


Figure 5.10 Screenshot of question after navigating the web site in Study 3

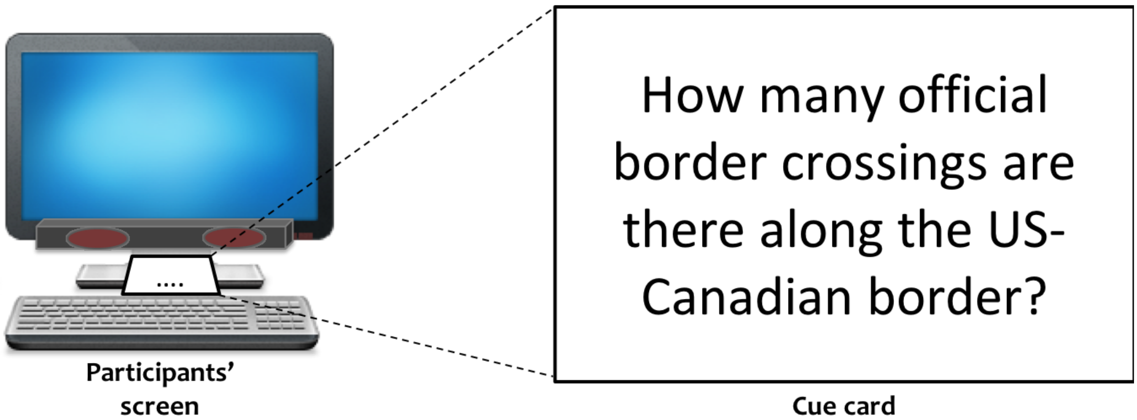


Figure 5.11 Position of cue card in Study 3

Two online post-study questionnaires were developed, which measured participants' opinions on ease of use, ease of remembering and ease of learning on the menus. In the first post-study questionnaire (see Appendix N), they were asked to rate the menus they had experienced in the tasks. In the second post-study questionnaire (see Appendix O), they were shown other menu combinations which they had not experienced and were asked to browse the experimental web site using all menu combinations and then rate all four menu combinations (Unified-Visible, Unified-Dynamic, Fragmented-Visible and Fragmented-Dynamic).

5.3.1.4 Procedure

Participants were asked to complete an ARQ (see Appendix C) before they came for their experiment session in the lab. This was done in order to screen for dyslexia symptoms and categorise them into the different participant groups. The experiment session was completed in a quiet room where they were given a brief explanation about the purpose of the study and the tasks they are going to undertake. Then, they were given a chance to ask any questions, before completing Section A of the informed consent form (see Appendix D) and the demographic questionnaire, covering their use of the Web and their personal details (see Appendix E).

Participant was asked to sit comfortably and adjust the height of the fixed chair and the gap between their body and the desk. The monitor screen was then adjusted so the viewing distance from the participant to the screen was approximately 60cm. The participant was asked to minimise head and body movement during the experiment to optimise the eye tracking equipment.

Participants were given a demonstration of how the Experiment Center™ application calibrates and validates eye gaze. They were told that the calibration and validation would be repeated if it failed to calibrate their eye gaze on the first attempt. As majority of participants had participated in my previous studies, they did the calibration and validation on their own, monitored by me. After a successful calibration, the application automatically directed the participant to the Practice Task.

At the beginning of the Practice Task, they were given a question with three alternative responses. However, they could only read the question without being able to choose any possible answer. The application automatically directed them to the practice web site. They navigated around the web site in order to find the answer (see Figure 5.4). Once they found the answer, they closed the browser. The application then presented the question again with a radio button on each answer, then they were able to choose one possible answer.

Then, the participant was assigned one version of the experimental web site (either Figure 5.5, Figure 5.6, Figure 5.7 or Figure 5.8). They were asked to complete four questions for the Experimental tasks. The calibration and validation of their eye gaze was repeated at the beginning of each task. After successfully calibrating and validating their eye gaze, they read the question (see Figure 5.9), were directed to the web page and navigated to find the answer, closed the browser and finally chose their answer (see

Figure 5.10). This process was repeated four times with different questions using the same version of the web site. Appendix P shows the sequence of questions which participants had in all four tasks.

After completion of all four questions in the Experimental Tasks, the participant completed two post-study questionnaires. In the first post-study questionnaire, they were shown the version of the experimental web site they had used and were asked to rate the menu combinations (see Appendix N). Then, the participant was shown all four versions of the experimental web site (experimental web site with Unified-Visible menus as in Figure 5.5, Unified-Dynamic menus as in Figure 5.6, Fragmented-Visible menus as in Figure 5.7, or Fragmented-Dynamic menus as in Figure 5.8). They were asked to browse the experimental web site using all menus and then rate all four menu combinations (see Appendix O). Participants were debriefed and any questions they had were answered. They were asked to sign Section B of the informed consent form (see Appendix D) to indicate they were satisfied with all the answers given. Participants then received their gift voucher through their email.

5.3.2 Data Preparation

Using the SMI BeGaze™, AOIs were created for main menus, sub-menus and target link in each task. As mentioned earlier in Section 5.3.1.1, each AOI was associated with similar variables. Results for entry time, dwell time, first fixation duration, number of revisits, number of fixations and average fixation duration, first mouse click time, task completion time and navigation score were taken from data recorded with the SMI Experiment Center™. Participants' opinions were taken from both post-study questionnaires given after the experimental task.

Entry time was taken when participants' eye gaze entered the AOI. The number of fixations and fixation duration was calculated for the time between when participants' eye gaze entered the AOI and left the AOI. First fixation duration was taken from the duration for the first fixation in the AOI. Average fixation duration was then calculated by dividing the fixation duration and number of fixations. Dwell time was calculated as a total of duration for fixations and saccades that happened in the AOI. The number of revisits was recorded when the participants returned their gaze to the AOI. First mouse click time was taken when the participants first clicked on any menu item. Task completion time was taken from the time the participants started browsing the web site until they closed the browser after they found the answer for the question.

Results from this study were grouped into three categories: (1) eye gaze behaviour (entry time, dwell time, first fixation duration, number of revisits, number of fixations and average fixation duration), (2) navigation performance (first mouse click time, task completion time and navigation score) and (3) participants' opinions. All variables in eye gaze behaviour categories, first mouse click time and task completion time were averaged for all tasks.

For each question, participant eye gaze in all AOIs was visually inspected using SMI BeGaze™ in order to confirm that the eye tracker had correctly recorded their eye gaze. If their eye gaze was not accurate, due to excessive head movement or other failures, their data was excluded from the analysis. One initial participant had inaccurate eye gaze recorded outside AOIs. The participant was replaced with another participant matched for age, gender and their severity of dyslexia. Seven participants completed one of the tasks really fast and SMI was unable to capture the web page screenshot within the short period of time. Due to non-existence of the screenshot, SMI failed to map their recorded eye gaze onto the screenshot and visualise it for inspection. Therefore, I had to manually inspect their eye gaze from the recorded footage. Then, I counted their eye gaze from the recorded raw data based on Cartesian coordinates (x and y axis) of drawn AOIs.

In each AOI, data were averaged across all questions for each variable in the eye gaze behaviour measures and some variable in the navigation performance measures (first mouse click time and task completion time). The data was categorised into the four different combinations of the two variables: menus with Unified main menus and Visible sub-menus was abbreviated as 'UV', Unified main menus and Dynamic sub-menus was abbreviated as 'UD', Fragmented main menus and Visible sub-menus was abbreviated as 'FV', and Fragmented main menus and Dynamic sub-menus was abbreviated as 'FD'.

All data were entered into a MS Excel spreadsheet and then analysed using the SPSS statistical program. All data were visually analysed using histogram to check for normal distribution. Outliers for all variables in eye gaze behaviour measures and some variables in navigation performance measures (first mouse click time and task completion time) were adjusted using *winsorization* technique (DeCator, 2015). In this technique, outliers were adjusted if the values below or above than $Mean \pm 2SD$.

5.3.3 Results

The majority of data for eye gaze behaviour, navigation performance and participants' opinions were normally distributed and variances homogeneous. As mentioned in previous studies (Chapter 3 and Chapter 4) , non-normality and heterogenous data can additively influence the results of parametric test ANOVA. The data were transformed with square root function in an attempt to have better distribution of normality and variances homogenous data. After the transformation, there was an improvement in the distribution of normality and variances homogenous for eye gaze behaviour and navigation performance, but not with participants' opinions. Then, data for participants' opinions were transformed again using Log10. However, the distribution of normality and variances homogeneity worsened compared to the original data. According to Laerd Statistics (2015), there are no equivalent non-parametric tests for the parametric test of three-way ANOVA. Therefore, I decided to run the parametric test on both original and transformed data to see if there are any meaningful differences.

For both original and transformed data in eye gaze behaviour and navigation performance, significance results were not appreciably different with the ANOVA. I decided to report the results of this measures using the original data with the parametric test of ANOVA since the ANOVA is robust enough to handle the violations when the sample size is equal (Blanca et al., 2017; Laerd Statistics, 2015; Schmider et al., 2010). However for participants' opinions, different significance results were found for both original and transformed data. I decided to report the results of this measures using the original data with the parametric tests of ANOVA since original data had better distribution of normality.

Bonferroni-corrected post-hoc tests were conducted to investigate specific differences uncovered by the analyses. An alpha significance level of .05 was used in all statistical tests in this study. For effect size, partial eta squared of .01 is small, .06 is medium and .14 is large was used to determine the magnitude of the statistical results (Draper, 2018). Details for each test are discussed in next sections.

5.3.3.1 Eye Gaze Behaviour

A 2 x 2 x 2 ANOVA was used on all three AOIs, comprising two participant groups (non-dyslexic and dyslexic), two organisations of menus (unified and fragmented) and two visibilities of sub-menus (visible and dynamic). The dependent variables for all AOIs were entry time, dwell time, first fixation duration, number of revisits, number of fixations and average fixation duration.

a. Entry Time

Table 5.2 shows the descriptive statistics for the entry time to the three AOIs for the two participant groups.

For the Main Menus AOI, the entry times were normally distributed except for Dyslexic participants in the Fragmented-Dynamic condition, *p* = .02, as assessed by Shapiro-Wilk's test of normality. The assumption of homogeneity of variances was not met (*F* = 2.30, *p* = .04), as assessed by Levene's test for equality of variances.

For the Sub-Menus AOI, the entry times were normally distributed except for Non-Dyslexic participants in the Fragmented-Visible condition, *p* = .01, as assessed by Shapiro-Wilk's test of normality. The assumption of homogeneity of variances was not met (*F* = 2.61, *p* = .02), as assessed by Levene's test for equality of variances.

For the Target Link AOI, the entry times were normally distributed, as assessed by Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

The results of the ANOVA are summarised in Table 5.4.

Table 5.2 Descriptive statistics of entry time (ms) for both participant groups in all AOIs

	Main Menus AOI		**Sub-Menus AOI**		**Target Link AOI**	
	Non-Dyslexic	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**
	M **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**
UV	763.29 (631.66)	598.04 (404.64)	479.07 (405.82)	438.21 (203.74)	2207.93 (1303.20)	2525.98 (1525.58)
UD	719.49 (675.56)	656.85 (637.36)	1149.02 (665.62)	823.59 (627.59)	3473.09 (846.69)	3657.15 (15.46.44)
FV	809.26 (577.16)	1047.50 (743.48)	352.77 (210.02)	735.19 (583.06)	3947.90 (2388.85)	3781.11 (1425.88)
FD	634.23 (356.66)	1542.92 (1700.16)	1100.94 (675.14)	1352.03 (561.83)	2763.39 (971.19)	5398.11 (2326.81)

Note. Measures in milliseconds (ms) per AOI for all tasks.

For the Main Menus AOI there were no significant effects.

For the Sub-Menus AOI, there was significant main effect for Visibility of sub-menus with large effect size, F(1, 56) = 21.32, p < .001, η_p^2 = .28, where Visible sub-menus (*M* = 501.31, *SD* = 392.75) had a significantly faster entry time than Dynamic sub-menus (*M* = 1106.40, *SD* = 632.35).

For the Target Link AOI, there was significant main effect for Organisation of menus with moderate to large effect size, *F*(1, 56) = 6.11, *p* = .02, η_p^2 = .10, where Unified menus (*M* = 2966.03, *SD* = 1413.80) had significant faster entry time than Fragmented menus (*M* = 3972.63, *SD* = 2023.50).

b. Dwell Time

Table 5.3 shows the descriptive statistics for dwell time in the three AOIs for the two participant groups.

For the Main Menus AOI, the dwell times were normally distributed, as assessed by Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

For the Sub-Menus AOI, the dwell times were normally distributed except for Dyslexic participants in the Unified-Dynamic condition, *p* < .01 and Non-Dyslexic participants in the Unified-Dynamic condition, *p* < .01, as assessed by Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

For the Target Link AOI, the dwell times were normally distributed except for Dyslexic participants in the Unified-Dynamic condition, *p* = .03, as assessed by Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

The results of the ANOVA are summarised in Table 5.5.

Table 5.3 Descriptive statistics of dwell time (ms) for both participant groups in all AOIs

	Main Menus AOI		**Sub-Menus AOI**		**Target Link AOI**	
	Non-Dyslexic	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**
	M **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**
UV	2434.06 (1125.35)	2385.88 (601.04)	1897.45 (424.53)	3174.42 (1102.47)	878.08 (121.30)	1090.28 (301.30)
UD	1976.88 (675.31)	2144.80 (1180.47)	2481.75 (1155.39)	3103.87 (902.36)	917.319 (297.96)	831.99 (333.54)
FV	1905.01 (892.66)	2260.16 (969.54)	3256.97 (1298.75)	3403.57 (1292.06)	1014.19 (391.09)	1039.14 (382.04)
FD	2016.49 (843.41)	2778.21 (739.37)	2690.53 (1160.42)	3029.18 (1618.16)	881.86 (488.96)	1020.91 (290.71)

Note. Measures in milliseconds (ms) per AOI for all tasks.

For the Main Menus AOI there were no significant effects.

For the Sub-Menus AOI, there was significant main effect for Group with moderate effect size, $F(1, 56) = 4.19$, $p < .05$, $\eta_p^2 = .07$. Non-Dyslexic participants had significantly shorter dwell times (M = 2581.68, SD = 1127.48) than Dyslexic participants (M = 3177.76, SD = 1202.84).

For the Target Link AOI there were no significant effects.

Table 5.4 Summary of three-way ANOVA on entry time for all AOIs

	Effect	Main Menus AOI			Sub-Menus AOI			Target Link AOI		
		F	***p***	η_p^2	***F***	***p***	η_p^2	***F***	***p***	η_p^2
Main	Group	$F(1, 56) = 1.27$	ns	.02	$F(1, 56) = 0.26$	ns	.01	$F(1, 56) = 3.32$	ns	.06
	Organisation	$F(1, 56) = 2.52$	ns	.04	$F(1, 56) = 1.54$	ns	.03	$F(1, 56) = 6.11$	.02	.10
	Visibility	$F(1, 56) = 0.17$	ns	.00	$F(1, 56) = 21.32$	< .001	.28	$F(1, 56) = 3.02$	ns	.05
Two-way interaction	Group × Organisation	$F(1, 56) = 2.84$	ns	.05	$F(1, 56) = 3.64$	ns	.06	$F(1, 56) = 1.46$	ns	.03
	Group × Visibility	$F(1, 56) = 0.90$	ns	.02	$F(1, 56) = 0.63$	ns	.01	$F(1, 56) = 2.68$	ns	.05
	Organisation × Visibility	$F(1, 56) = 0.14$	ns	.00	$F(1, 56) = 0.35$	ns	.01	$F(1, 56) = 1.45$	ns	.03
Three-way interaction	Group × Organisation × Visibility	$F(1, 56) = 0.48$	ns	.01	$F(1, 56) = 0.09$	ns	.00	$F(1, 56) = 3.25$	ns	.06

Table 5.5 Summary of three-way ANOVA on dwell time for all AOIs

	Effect	Main Menus AOI			Sub-Menus AOI			Target Link AOI		
		F	***p***	η_p^2	***F***	***p***	η_p^2	***F***	***p***	η_p^2
Main	Group	$F(1, 56) = 1.89$	ns	.03	$F(1, 56) = 4.19$	< .05	.07	$F(1, 56) = 0.73$	ns	.01
	Organisation	$F(1, 56) = 0.00$	ns	.00	$F(1, 56) = 2.19$	ns	.04	$F(1, 56) = 0.49$	ns	.01
	Visibility	$F(1, 56) = 0.01$	ns	.00	$F(1, 56) = 0.13$	ns	.00	$F(1, 56) = 1.18$	ns	.02
Two-way interaction	Group × Organisation	$F(1, 56) = 1.23$	ns	.02	$F(1, 56) = 1.47$	ns	.03	$F(1, 56) = 0.01$	ns	.00
	Group × Visibility	$F(1, 56) = 0.48$	ns	.01	$F(1, 56) = 0.16$	ns	.00	$F(1, 56) = 0.29$	ns	.01
	Organisation × Visibility	$F(1, 56) = 2.18$	ns	.04	$F(1, 56) = 1.56$	ns	.03	$F(1, 56) = 0.04$	ns	.00
Three-way interaction	Group × Organisation × Visibility	$F(1, 56) = 0.05$	ns	.00	$F(1, 56) = 0.53$	ns	.01	$F(1, 56) = 1.46$	ns	.03

c. First Fixation Duration

Table 5.6 shows the descriptive statistics for the first fixation duration in the three AOIs for the two participant groups.

For all AOIs, the first fixation durations were normally distributed, as assessed by Shapiro-Wilk's test of normality. There were no violations of homogeneity of variances, as assessed by Levene's test for equality of variances.

The results of the ANOVA are summarised in Table 5.8.

Table 5.6 Descriptive statistics of first fixation duration (ms) for both participant groups in all AOIs

	Main Menus AOI		**Sub-Menus AOI**		**Target Link AOI**	
	Non-Dyslexic	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**
	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**
UV	159.74 (57.76)	160.22 (54.28)	212.58 (110.10)	218.26 (83.52)	238.63 (92.94)	263.75 (52.41)
UD	165.00 (31.70)	193.08 (70.91)	210.18 (34.65)	227.66 (61.34)	247.61 (85.01)	274.33 (39.59)
FV	178.46 (50.56)	185.97 (63.43)	232.58 (58.18)	289.33 (71.76)	241.20 (57.99)	279.05 (132.87)
FD	183.31 (90.98)	201.63 (79.52)	258.58 (84.85)	194.05 (49.32)	306.59 (107.74)	196.32 (52.82)

Note. Measures in milliseconds (ms) per AOI for all tasks.

No significant effect was found on first fixation duration for all AOIs in this study.

d. Number of Revisits

Table 5.7 shows the descriptive statistics for the number of revisits in the three AOIs for the two participant groups.

For the Main Menus AOI, the number of revisits were normally distributed except for Dyslexic participants in the Unified-Visible condition, $p < .01$ and Non-Dyslexic participants in the Fragmented-Visible condition, $p = .04$, as assessed by Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

For the Sub-Menus AOI, the number of revisits were normally distributed except for Non-Dyslexic participants in the Fragmented-Visible condition, $p = .02$, as assessed by

Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

For the Target Link AOI, the number of revisits were normally distributed except for Non-Dyslexic participants in the Fragmented-Visible condition, *p* = .04, as assessed by Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

The results of the ANOVA are summarised in Table 5.9.

Table 5.7 Descriptive statistics of number of revisits for both participant groups in all AOIs

	Main Menus AOI		**Sub-Menus AOI**		**Target Link AOI**	
	Non-Dyslexic	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**
	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**
UV	3.31 (1.60)	3.53 (1.09)	1.59 (0.63)	2.48 (1.05)	0.88 (0.42)	1.50 (0.58)
UD	2.15 (1.11)	2.88 (1.36)	2.17 (1.66)	2.59 (0.86)	0.88 (0.60)	0.94 (0.55)
FV	3.10 (1.20)	2.81 (1.56)	2.97 (1.11)	2.75 (1.60)	1.17 (0.69)	0.97 (0.59)
FD	1.75 (0.40)	2.84 (0.84)	1.68 (1.21)	3.00 (1.90)	0.77 (0.37)	1.15 (0.47)

Note. Measures in count per AOI for all tasks.

For the Main Menus AOI, there was a significant main effect of Visibility with moderate to large effect size, $F(1, 56) = 6.86$, $p = .01$, $\eta_p^2 = .11$. Menus with Dynamic sub-menus ($M = 2.40$, $SD = 1.06$) had significantly lower number of revisits than menus with Visible sub-menus ($M = 3.19$, $SD = 1.34$).

For the Sub-Menus AOI there were no significant effects.

For the Target Link AOI, there was a significant three-way interaction of Group, Organisation and Visibility with moderate effect size, $F(1, 56) = 4.39$, $p = .04$, $\eta_p^2 = .07$. Figure 5.12 illustrates this interaction. Further post-hoc analysis indicated that this three-way interaction was not due to the presence of any significant simple two-way interactions, which mean the differences happened due to random sampling. However, there is a marginal trend with large effect size for a two-way interaction of Group and Organisation for Visible sub-menus with moderate to large effect size, $F(1, 28) = 4.05$, $p = .05$, $\eta_p^2 = .13$.

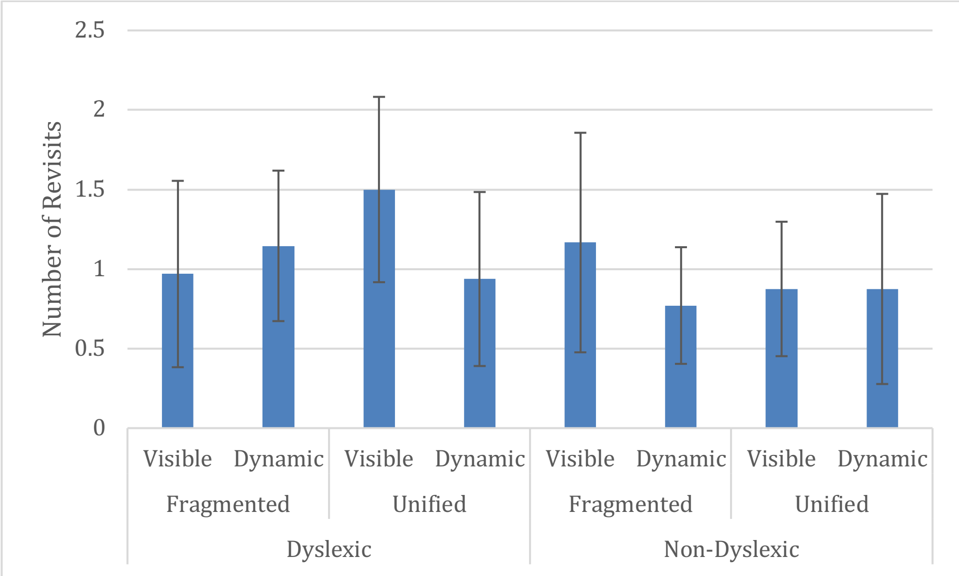


Figure 5.12 Mean number of revisits for the two participant groups

From Figure 5.12, it can be seen that Dyslexic participants revisited the target link in the Unified-Visible condition (*M* = 1.50, *SD* = 0.58) more than in the Unified-Dynamic condition (*M* = 0.94, *SD* = 0.55) and Non-Dyslexic participants revisited the target link in the Fragmented-Visible condition (*M* = 1.17, *SD* = 0.69) more than in the Fragmented-Dynamic condition (*M* = 0.77, *SD* = 0.37).

Table 5.8 Summary of three-way ANOVA on first fixation duration for all AOIs

	Effect	Main Menus AOI			Sub-Menus AOI			Target Link AOI		
		F	***p***	η_p^2	***F***	***p***	η_p^2	***F***	***p***	η_p^2
Main	Group	$F(1, 56) = 0.71$	ns	.01	$F(1, 56) = 0.05$	ns	.00	$F(1, 56) = 0.06$	ns	.00
	Organisation	$F(1, 56) = 1.22$	ns	.02	$F(1, 56) = 2.12$	ns	.04	$F(1, 56) = 0.00$	ns	.00
	Visibility	$F(1, 56) = 0.83$	ns	.02	$F(1, 56) = 0.74$	ns	.01	$F(1, 56) = 0.00$	ns	.00
Two-way interaction	Group × Organisation	$F(1, 56) = 0.00$	ns	.00	$F(1, 56) = 0.18$	ns	.00	$F(1, 56) = 2.23$	ns	.04
	Group × Visibility	$F(1, 56) = 0.36$	ns	.01	$F(1, 56) = 2.27$	ns	.04	$F(1, 56) = 3.10$	ns	.05
	Organisation × Visibility	$F(1, 56) = 0.08$	ns	.00	$F(1, 56) = 1.10$	ns	.02	$F(1, 56) = 0.20$	ns	.00
Three-way interaction	Group × Organisation × Visibility	$F(1, 56) = 0.07$	ns	.00	$F(1, 56) = 3.36$	ns	.06	$F(1, 56) = 3.23$	ns	.06

Table 5.9 Summary of three-way ANOVA on number of revisits for all AOIs

	Effect	Main Menus AOI			Sub-Menus AOI			Target Link AOI		
		F	***p***	η_p^2	***F***	***p***	η_p^2	***F***	***p***	η_p^2
Main	Group	$F(1, 56) = 2.12$	ns	.04	$F(1, 56) = 3.38$	ns	.06	$F(1, 56) = 2.54$	ns	.04
	Organisation	$F(1, 56) = 1.27$	ns	.02	$F(1, 56) = 1.41$	ns	.03	$F(1, 56) = 0.06$	ns	.00
	Visibility	$F(1, 56) = 6.86$	.01	.11	$F(1, 56) = 0.07$	ns	.00	$F(1, 56) = 2.08$	ns	.04
Two-way interaction	Group × Organisation	$F(1, 56) = 0.02$	ns	.00	$F(1, 56) = 0.03$	ns	.00	$F(1, 56) = 0.89$	ns	.02
	Group × Visibility	$F(1, 56) = 2.49$	ns	.04	$F(1, 56) = 0.68$	ns	.01	$F(1, 56) = 0.00$	ns	.00
	Organisation × Visibility	$F(1, 56) = 0.17$	ns	.00	$F(1, 56) = 1.73$	ns	.03	$F(1, 56) = 0.40$	ns	.01
Three-way interaction	Group × Organisation × Visibility	$F(1, 56) = 0.53$	ns	.01	$F(1, 56) = 2.31$	ns	.04	$F(1, 56) = 4.39$	.04	.07

e. Number of Fixations

Table 5.10 shows the descriptive statistics for the number of fixations in the three AOIs for the two participant groups.

For the Main Menus AOI, the number of fixations were normally distributed except for Dyslexic participants in the Unified-Visible condition, p = .03, as assessed by Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

For the Sub-Menus AOI, the number of fixations were normally distributed except for Dyslexic participants in the Unified-Dynamic condition, p = .04 and Non-Dyslexic participants in the Fragmented-Visible condition, p = .03, as assessed by Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

For the Target Link AOI, the number of fixations were normally distributed, as assessed by Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

The results of the ANOVA are summarised in Table 5.12.

Table 5.10 Descriptive statistics of number of fixations for both participant groups in all AOIs

	Main Menus AOI		**Sub-Menus AOI**		**Target Link AOI**	
	Non-Dyslexic	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**
	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**
UV	10.94 (3.54)	11.36 (2.75)	8.03 (1.92)	12.44 (4.07)	3.54 (0.47)	4.31 (1.21)
UD	9.06 (2.11)	8.66 (4.40)	9.24 (3.71)	11.49 (2.96)	3.38 (0.78)	3.00 (1.38)
FV	7.69 (3.02)	9.71 (3.61)	12.72 (4.35)	12.91 (4.59)	3.79 (1.13)	4.28 (1.76)
FD	8.80 (3.38)	12.03 (2.44)	9.76 (4.00)	12.40 (5.25)	3.33 (1.73)	3.93 (1.06)

Note. Measures count per AOI for all tasks.

For the Main Menus AOI, there was a significant two-way interaction for Organisation and Visibility with moderate to large effect size, $F(1, 56) = 6.17$, $p = .02$, $\eta_p^2 = .10$. Figure 5.13 illustrates this interaction. Post-hoc analysis indicated that Unified menus produced a significantly higher number of fixations when combined with Visible sub-menus than when combined with Dynamic sub-menus (p < .05). In addition, Unified menus had a significantly higher number of fixations than Fragmented menus when combined with Visible sub-menus (p = .04).

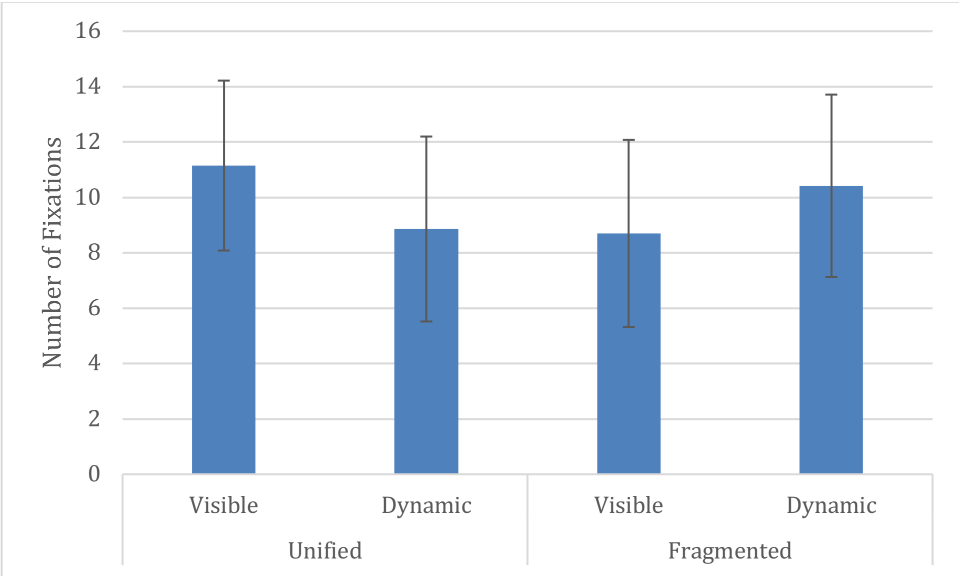


Figure 5.13 Mean of number of fixations for all combinations of main menus and sub-menus

For the Sub-Menus AOI, there was significant main effect for Group with moderate to large effect size, $F(1, 56) = 5.69$, $p = .02$, $\eta_p^2 = .09$. Non-Dyslexic participants (M = 9.94, SD =3.86) had a significantly lower number of fixations than Dyslexic participants (M = 12.31, SD = 4.12).

For the Target Link AOI there were no significant effects.

f. Average Fixation Duration

Table 5.11 shows the descriptive statistics for the average fixation duration in the three AOIs for the two participant groups.

For the Main Menus AOI, the average fixation durations were normally distributed, as assessed by Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

For the Sub-Menus AOI, the average fixation durations were normally distributed except for Non-Dyslexic participants in the Unified-Visible condition, p = .02, as assessed by

Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

For the Target Link AOI, the average fixation durations were normally distributed, as assessed by Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

The results of the ANOVAs are summarised in Table 5.13.

Table 5.11 Descriptive statistics of average fixation duration (ms) for both participant groups in all AOIs

	Main Menus AOI		**Sub-Menus AOI**		**Target Link AOI**	
	Non-Dyslexic	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**
	M **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**
UV	200.31 (53.42)	180.46 (23.45)	223.97 (41.69)	245.94 (36.27)	245.75 (81.64)	255.77 (53.16)
UD	198.69 (46.71)	232.84 (58.35)	238.55 (61.25)	237.11 (41.18)	270.99 (62.39)	273.01 (47.32)
FV	208.60 (41.40)	206.79 (35.74)	235.11 (40.04)	246.34 (54.42)	263.71 (59.51)	262.37 (92.08)
FD	214.60 (43.23)	208.60 (41.90)	244.08 (40.48)	201.91 (45.87)	265.98 (49.95)	245.03 (74.54)

Note. Measures in milliseconds (ms) per AOI for all tasks.

No significant effect was found on first fixation duration for all AOIs in this study.

Table 5.12 Summary of three-way ANOVA on number of fixations for all AOIs

	Effect	Main Menus AOI			Sub-Menus AOI			Target Link AOI		
		F	***p***	η_p^2	***F***	***p***	η_p^2	***F***	***p***	η_p^2
Main	Group	$F(1, 56) = 2.67$	ns	.05	$F(1, 56) = 5.69$	.02	.09	$F(1, 56) = 1.38$	ns	.02
	Organisation	$F(1, 56) = 0.31$	ns	.01	$F(1, 56) = 2.75$	ns	.05	$F(1, 56) = 0.77$	ns	.01
	Visibility	$F(1, 56) = 0.13$	ns	.00	$F(1, 56) = 0.65$	ns	.01	$F(1, 56) = 3.31$	ns	.06
Two-way interaction	Group × Organisation	$F(1, 56) = 2.62$	ns	.05	$F(1, 56) = 0.93$	ns	.02	$F(1, 56) = 0.30$	ns	.01
	Group × Visibility	$F(1, 56) = 0.01$	ns	.00	$F(1, 56) = 0.01$	ns	.00	$F(1, 56) = 0.68$	ns	.01
	Organisation × Visibility	$F(1, 56) = 6.17$	.02	.10	$F(1, 56) = 0.88$	ns	.02	$F(1, 56) = 0.28$	ns	.01
Three-way interaction	Group × Organisation × Visibility	$F(1, 56) = 0.40$	ns	.01	$F(1, 56) = 1.34$	ns	.02	$F(1, 56) = 0.99$	ns	.02

Table 5.13 Summary of three-way ANOVA on average fixation duration for all AOIs

	Effect	Main Menus AOI			Sub-Menus AOI			Target Link AOI		
		F	***p***	η_p^2	***F***	***p***	η_p^2	***F***	***p***	η_p^2
Main	Group	$F(1, 56) = 0.02$	ns	.00	$F(1, 56) = 0.05$	ns	.00	$F(1, 56) = 0.02$	ns	.00
	Organisation	$F(1, 56) = 0.35$	ns	.01	$F(1, 56) = 0.16$	ns	.00	$F(1, 56) = 0.02$	ns	.00
	Visibility	$F(1, 56) = 1.76$	ns	.03	$F(1, 56) = 0.42$	ns	.01	$F(1, 56) = 0.17$	ns	.00
Two-way interaction	Group × Organisation	$F(1, 56) = 0.25$	ns	.00	$F(1, 56) = 1.26$	ns	.02	$F(1, 56) = 0.26$	ns	.01
	Group × Visibility	$F(1, 56) = 1.27$	ns	.02	$F(1, 56) = 2.81$	ns	.05	$F(1, 56) = 0.17$	ns	.00
	Organisation × Visibility	$F(1, 56) = 0.95$	ns	.02	$F(1, 56) = 0.81$	ns	.01	$F(1, 56) = 0.74$	ns	.01
Three-way interaction	Group × Organisation × Visibility	$F(1, 56) = 1.74$	ns	.03	$F(1, 56) = 0.43$	ns	.01	$F(1, 56) = 0.03$	ns	.00

5.3.3.2 Navigation Performance

a. First Mouse Click Time

Table 5.14 shows the descriptive statistics for the first mouse click time for the two participant groups. First mouse click times were normally distributed, as assessed by Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

The results of the ANOVA are summarised in Table 5.15.

Table 5.14 Descriptive statistics of first mouse click time for both participant groups

	Non-Dyslexic		**Dyslexic**	
	M	***SD***	***M***	***SD***
UV	4698.67	1631.68	6330.65	2268.77
UD	5971.63	1481.48	7212.43	2330.47
FV	6924.17	3054.70	7613.68	3939.47
FD	4654.86	838.69	8469.60	3995.98

Note. Measures in milliseconds (ms) for all tasks.

Table 5.15 Summary of three-way ANOVA on first mouse click time

	Effect	***F***	***p***	η_p^2
Main	Group	$F(1, 56) = 7.64$	<.01	.12
	Organisation	$F(1, 56) = 1.67$	ns	.03
	Visibility	$F(1, 56) = 0.07$	ns	.00
Two-way interaction	Group × Organisation	$F(1, 56) = 0.37$	ns	.01
	Group × Visibility	$F(1, 56) = 1.05$	ns	.02
	Organisation × Visibility	$F(1, 56) = 1.79$	ns	.03
Three-way interaction	Group × Organisation × Visibility	$F(1, 56) = 1.74$	ns	.03

There was a significant main effect for Group with moderate to large effect size, $F(1, 56) = 7.64$, $p < .01$, $\eta_p^2 = .12$. Non-Dyslexic participants ($M = 5562.33$, $SD = 2071.11$) were significantly faster than Dyslexic participants ($M = 7406.59$, $SD = 3179.60$).

b. Task Completion Time

Table 5.16 shows the descriptive statistics for the task completion times for the two participant groups. Task completion times were normally distributed, as assessed by Shapiro-Wilk's test of normality. There was no violation of homogeneity of variances, as assessed by Levene's test for equality of variances.

The results of the ANOVA are summarised in Table 5.17.

Table 5.16 Descriptive statistics of task completion time for both participant groups

	Non-Dyslexic		Dyslexic	
	M	***SD***	***M***	***SD***
UV	23293.16	7461.54	32919.27	4897.27
UD	29518.09	10085.59	27223.19	7987.87
FV	31042.30	15112.97	36484.22	14852.24
FD	23249.80	7897.05	35253.73	12582.71

Note. Measures in milliseconds (ms) for all tasks.

Table 5.17 Summary of three-way ANOVA on task completion times

	Effect	***F***	***p***	η_p^2
Main	Group	*F*(1, 56) = 5.34	.02	.09
	Organisation	*F*(1, 56) = 1.50	ns	.03
	Visibility	*F*(1, 56) = 0.63	ns	.01
Two-way interaction	Group × Organisation	*F*(1, 56) = 0.90	ns	.02
	Group × Visibility	*F*(1, 56) = 0.25	ns	.00
	Organisation × Visibility	*F*(1, 56) = 0.80	ns	.01
Three-way interaction	Group × Organisation × Visibility	*F*(1, 56) = 2.99	ns	.05

There was a significant main effect for Group with moderate to large effect size, $F(1, 56) = 5.34$, $p = .02$, $\eta_p^2 = .09$. Non-Dyslexic participants (*M* = 26775.84, *SD* = 10685.22) were significantly faster than Dyslexic participants (*M* = 32970.10, *SD* = 10883.05).

c. Navigation Scores

Participants answered a total of four questions across the four experimental tasks. The median total navigation score for non-dyslexic participants (*Mdn* = 4.00, *IQR* = 0.00) was not different from dyslexic participants (*Mdn* = 4.00, *IQR* = 0.00). The navigation scores were not normally distributed for both non-dyslexic and dyslexic participants (both $p < .001$), as assessed by Shapiro-Wilk's test for normality. Kruskal-Wallis test has shown that median of the score were not statistically significant between two participant groups, $X^2(1) = 0.00$, $p = 1.00$.

5.3.3.3 Participants' Opinions

a. Ease of Use, Ease of Remembering, Ease of Learning of The Menus Experienced in the Tasks

Table 5.18 shows the descriptive statistics for the ratings on ease of use, ease of remembering and ease of learning for the two participant groups for the combination of menu variables they experienced in the tasks. Note that lower ratings indicate greater difficulty while higher ratings indicate greater ease.

Table 5.18 Descriptive statistics of ratings on ease of use, ease of remembering and ease of learning for participants' assigned menus combination for both participant groups

	Ease of Use		Ease of Remembering		Ease of Learning	
	Non-Dyslexic	Dyslexic	Non-Dyslexic	Dyslexic	Non-Dyslexic	Dyslexic
	M **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**
UV	6.12 (1.46)	5.75 (2.05)	5.75 (1.17)	5.50 (1.77)	6.25 (0.89)	5.75 (1.58)
UD	6.25 (1.75)	6.50 (0.54)	5.75 (0.71)	5.75 (0.46)	6.12 (1.13)	6.00 (1.31)
FV	4.75 (1.83)	5.87 (0.84)	4.37 (1.41)	5.12 (0.84)	5.12 (1.46)	5.12 (1.55)
FD	5.75 (1.04)	5.00 (1.31)	4.87 (1.13)	5.37 (1.41)	5.37 (1.19)	5.37 (1.60)

Ratings on ease of use for dyslexic participants were not normally distributed in the Fragmented-Visible condition ($p < .01$), the Unified-Visible condition ($p < .01$) and the Unified-Dynamic condition ($p < .01$), as assessed by Shapiro-Wilk's test of normality. For non-dyslexic participants, rating for ease of use were normally distributed except for the Unified-Visible condition ($p < .01$) and the Unified-Dynamic condition ($p < .001$), as assessed by Shapiro-Wilk's test of normality. The assumptions of homogeneity of variances were met, as assessed Levene's test for equality of variances.

Ratings on ease of remembering for dyslexic participants were normally distributed except for the Unified-Dynamic condition ($p < .001$), as assessed by Shapiro-Wilk's test of normality. For non-dyslexic participants, rating for ease of remembering were normally distributed except for the Fragmented-Visible condition ($p = .02$), as assessed by Shapiro-Wilk's test of normality. The assumptions of homogeneity of variances were not met ($F = 2.79$, $p = .02$), as assessed Levene's test for equality of variances.

Ratings on ease of learning for dyslexic participants were normally distributed except for the Unified-Visible condition ($p < .001$) and the Unified-Dynamic condition ($p < .01$), as assessed by Shapiro-Wilk's test of normality. For non-dyslexic participants, rating for ease of learning were normally distributed except for the Unified-Visible condition ($p = .02$) and the Unified-Dynamic condition ($p = .04$), as assessed by Shapiro-Wilk's test of normality. The assumptions of homogeneity of variances were met, as assessed Levene's test for equality of variances.

The results of the ANOVAs on all of the ratings are summarised in Table 5.19.

For ratings on ease of use, there was a significant main effect for Organisation of menus with moderate to large effect size, $F(1, 56) = 5.11$, $p = .03$, $\eta_p^2 = .08$. Unified menus ($M = 6.16$, $SD = 1.50$) were rated as significantly easier to use than Fragmented menus ($M = 5.33$, $SD = 1.34$).

For ratings on ease of remembering, there was a significant main effect for Organisation of menus with moderate to large effect size, $F(1, 56) = 6.46$, $p = .01$, $\eta_p^2 = .10$. Unified menus ($M = 5.69$, $SD = 1.09$) were rated as significantly easier to remember than Fragmented menus ($M = 4.94$, $SD = 1.22$).

For ratings on ease of learning, there was a significant main effect for Organisation of menus with moderate to large effect size, $F(1, 56) = 5.29$, $p = .03$, $\eta_p^2 = .09$. Unified menus ($M = 6.03$, $SD = 1.20$) were rated as significantly easier to learn than Fragmented menus ($M = 5.25$, $SD = 1.39$).

Table 5.19 Summary of three-way ANOVAs on ratings of ease of use, ease of remembering and ease of learning for participants' assigned menu combination.

	Effect	Ease of Use			Ease of Remembering			Ease of Learning		
		F	***p***	η_p^2	***F***	***p***	η_p^2	***F***	***p***	η_p^2
Main	Group	$F(1, 56) = 0.03$	ns	.00	$F(1, 56) = 0.72$	ns	.01	$F(1, 56) = 0.21$	ns	.00
	Organisation	$F(1, 56) = 5.11$	.03	.08	$F(1, 56) = 6.46$	.01	.10	$F(1, 56) = 5.29$	.03	.09
	Visibility	$F(1, 56) = 0.48$	ns	.01	$F(1, 56) = 0.72$	ns	.01	$F(1, 56) = 0.21$	ns	.00
Two-way interaction	Group × Organisation	$F(1, 56) = 0.12$	ns	.00	$F(1, 56) = 1.62$	ns	.03	$F(1, 56) = 0.21$	ns	.00
	Group × Visibility	$F(1, 56) = 0.76$	ns	.01	$F(1, 56) = 0.00$	ns	.00	$F(1, 56) = 0.08$	ns	.00
	Organisation × Visibility	$F(1, 56) = 0.27$	ns	.01	$F(1, 56) = 0.18$	ns	.00	$F(1, 56) = 0.08$	ns	.00
Three-way interaction	Group × Organisation × Visibility	$F(1, 56) = 3.02$	ns	.05	$F(1, 56) = 0.18$	ns	.00	$F(1, 56) = 0.08$	ns	.00

b. Ease of Use, Ease of Remembering, Ease of Learning of All the Menu Combinations

Table 5.20 shows the descriptive statistics for the ratings on ease of use, ease of remembering and ease of learning for the two participant groups for all the menu combinations, the one the participants experienced plus the others which they were shown at the end of the experiment. Note that lower ratings indicate greater difficulty while higher ratings indicate greater ease. As mentioned earlier, majority of data violated the assumption of normality.

Table 5.20 Descriptive statistics of ratings on ease of use, ease of remembering and ease of learning for all menu combinations for both participant groups

	Ease of Use		**Ease of Remembering**		**Ease of Learning**	
	Non-Dyslexic	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**
	M **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**
UV	5.66 (1.56)	6.06 (1.22)	5.44 (1.80)	5.84 (1.30)	5.91 (1.59)	6.25 (1.02)
UD	6.06 (1.13)	6.03 (1.09)	5.78 (1.10)	5.41 (1.04)	6.22 (0.94)	6.00 (1.05)
FV	4.31 (1.58)	4.50 (1.39)	4.03 (1.77)	5.12 (1.36)	4.59 (1.78)	4.94 (1.29)
FD	4.31 (1.47)	4.03 (1.64)	4.16 (1.55)	4.03 (1.60)	4.66 (1.18)	4.66 (1.72)

Ratings of ease of use violated the assumption or normality except for dyslexic participants on the Fragmented-Visible condition (*p* = .13) and dyslexic (*p* = .09), as assessed by Shapiro-Wilk's test of normality. There were no violations of homogeneity of variances, as assessed by Levene's test for equality of variances. Mauchly's test of sphericity indicated that the assumption of sphericity was met.

Ratings of ease of remembering violated the assumption or normality except for non-dyslexic participants on the Fragmented-Dynamic condition (*p* = .08), as assessed by Shapiro-Wilk's test of normality. There were no violations of homogeneity of variances except for the Fragmented-Visible (*F* = 5.06, *p* = .03) and Unified-Visible (*F* = 4.23, *p* = .04) conditions, as assessed by Levene's test for equality of variances. Mauchly's test of sphericity indicated that the assumption of sphericity was met.

Ratings of ease of learning violated the assumption or normality, as assessed by Shapiro-Wilk's test of normality. There were no violations of homogeneity of variances except for the Fragmented-Dynamic condition (F = 5.90, p = .02), as assessed by Levene's test for equality of variances. Mauchly's test of sphericity indicated that the assumption of sphericity was met.

The results of the ANOVAs on all of the ratings are summarised in Table 5.21. From the table, interaction effects marked in red were significant when the ANOVAs were conducted with original data but not significant when the ANOVAs were conducted with transformed data. In addition, the effect sizes for the two interaction effects were medium. Since both original and transformed data were not normally distributed and variances heterogenous, the results of the two significant interaction effects are discarded and not discussed in this study.

For ratings of ease of use, there was a significant main effect for Organisation of menus with large effect size, $F(1, 62) = 78.97$, $p < .001$, $\eta_p^2 = .56$. Unified menus (M = 5.95, SD = 0.87) were rated significantly easier to use than Fragmented menus (M = 4.29, SD = 1.26).

For ratings of ease of remembering, there was a significant main effect for Organisation of menus with large effect size, $F(1, 62) = 62.98$, $p < .001$, $\eta_p^2 = .50$. Unified menus (M = 5.62, SD = 0.87) were rated as significantly easier to remember than Fragmented menus (M = 4.34, SD = 1.15).

For ratings on ease of learning, there was a significant main effect for Organisation of menus with large effect size, $F(1, 62) = 80.17$, $p < .001$, $\eta_p^2 = .56$. Unified menus (M = 6.09, SD = 0.87) were rated as significantly easier to learn than Fragmented menus (M = 4.69, SD = 1.20).

Table 5.21 Summary of three-way mixed ANOVAs on ratings of ease of use, ease of remembering and ease of learning for all menus combinations

	Effect	Ease of Use			Ease of Remembering			Ease of Learning		
		F	***p***	η_p^2	***F***	***p***	η_p^2	***F***	***p***	η_p^2
Main	Group	$F(1, 62) = 0.13$	ns	.00	$F(1, 62) = 1.64$	ns	.03	$F(1, 62) = 0.20$	ns	.00
	Organisation	$F(1, 62) = 78.97$	< .001	.56	$F(1, 62) = 62.98$	< .001	.50	$F(1, 62) = 80.17$	< .001	.56
	Visibility	$F(1, 62) = 0.02$	ns	.00	$F(1, 62) = 1.25$	ns	.02	$F(1, 62) = 0.10$	ns	.00
Two-way interaction	Group × Organisation	$F(1, 62) = 0.39$	ns	.01	$F(1, 62) = 2.11$	ns	.03	$F(1, 62) = 0.04$	ns	.00
	Group × Visibility	$F(1, 62) = 1.39$	ns	.02	$F(1, 62) = 4.42$	.04	.07	$F(1, 62) = 1.63$	ns	.03
	Organisation × Visibility	$F(1, 62) = 4.07$	.05	.06	$F(1, 62) = 3.64$	ns	.06	$F(1, 62) = 1.17$	ns	.02
Three-way interaction	Group × Organisation × Visibility	$F(1, 62) = 0.01$	ns	.00	$F(1, 62) = 0.91$	ns	.01	$F(1, 62) = 0.13$	ns	.00

Note. Results marked in red indicates the significant effects were discarded.

c. Consistency on the Given Ratings Between Experienced and Non-Experienced Menu Combinations

In the post-study questionnaire, participants were asked to rate the menu combination they had experienced in the study by undertaking tasks with it (T1). Then, they were asked to rate again the menu combinations in comparison with the other menu combinations used in this study and experienced by other participants (T2). Table 5.22 shows the descriptive statistics for the ratings on ease of use, ease of remembering and ease of learning given for the menu combinations for both occasions. Note that lower ratings of ease of use indicate greater difficulty while higher ratings indicate greater ease.

Table 5.22 Descriptive statistics of ratings on ease of use, ease of remembering and ease of learning of menu combinations for both occasions

	Ease of Use		**Ease of Remembering**		**Ease of Learning**	
	Non-Dyslexic	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**	**Non-Dyslexic**	**Dyslexic**
	M **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**	***M*** **(*SD*)**
T1	5.72 (1.59)	5.78 (1.36)	5.19 (1.23)	5.44 (1.19)	5.72 (1.22)	5.56 (1.48)
T2	5.50 (1.70)	5.66 (1.58)	5.06 (1.68)	5.69 (1.36)	5.47 (1.61)	5.75 (1.44)

For ratings of ease of use, the assumption of normality was not violated except for T1, $p < .01$. For ratings of ease of remembering, the assumption of normality was not violated except for T2, $p = .01$. For ratings of ease of learning, the assumption of normality was not violated except for T2, $p < .01$. All assumptions of normality were assessed by Shapiro-Wilk's test of normality.

From the paired samples t-tests, there were no significant changes found in participants' ratings for both occasions for ease of use ($t(63) = 0.98$, $p = .33$), ease of remembering ($t(63) = -0.35$, $p = .73$) and ease of learning ($t(63) = 0.16$, $p = .87$).

5.3.4 Discussion and Conclusion

This study investigated the effects of menu organisation and visibility on eye gaze behaviour, navigation performance and opinion measures on web navigation using eye tracking equipment. This study compared adult English native speakers with and without dyslexia. Participants were asked to complete an online dyslexia checklist (the ARQ) to categorise them as non-dyslexic and dyslexic.

Overall, this study found that organisation of menus had significant effects on number of fixations for main menus, on entry time for target link and participants' opinion. Visibility of sub-menus had significant effects on number of revisits for main menus and entry time to sub-menus. Participant group had significant effects dwell time and number of fixations for sub-menus, first mouse click time and task completion time in the navigation performance measures. Details for each of these aspects are discussed in the following sub-sections.

The term 'target link' used in this discussion section is referring to menu items that participants should click in order to access the web page that contains correct answer for the question asked in the experimental task.

5.3.4.1 Effects of Organisation of Menus

Guidelines for people with dyslexia recommend having clear and consistent menus and structure (de Santana et al., 2012; Evett & Brown, 2005; Friedman & Bryen, 2007; Zarach, 2002). However, there is no clear specification on how to make a clear and consistent menu for people with dyslexia. The guidelines for dyslexia also include other recommendations such as using non-transparent dynamic menus, making the menus visible all the time, using index page, customisable placement of menus, large menu, underline hyperlinks with clear label, legible font size for breadcrumb trail and alternative text on pictured menu items.

In a review of navigation studies with non-disabled users, Murano and Sander (2016) showed that users have better performance with menus placed horizontally at the top of the page. However for people with learning disabilities, placement of menus had no significant effect (Williams & Hennig, 2015a). As it has been recommended that the placement of menus have the ability to be customised (Friedman & Bryen, 2007), it is interesting to know if there are any significant effects of menu placement on people with dyslexia.

Previous studies have shown that a higher number of fixations reflects the attention given by participants to material (Farnsworth, 2018; Fitts, Jones, & Milton, 2005). The present study found that when menus were Unified, main menus had a higher number of fixations when combined with Visible sub-menus compared with Dynamic sub-menus. Participants probably gave full attention to the menus when they were organised in one place, and therefore fixated more as all the items in sub-menus were located in close proximity to the main menus. Not only that, when main menus were combined with Visible sub-menus, main menus that were Fragmented had a lower number of fixations than Unified main menus. This indicates that participants probably divided their attention when the menus were separated into different positions. Furthermore, participants may not have read all the items in the Fragmented menus.

This present study also found that participants noticed target links earlier when menus were Unified compared when they were Fragmented. It is easier to find the target link when all menu items are located at one single place, rather than scanning in two different areas. In addition, fragmented main menus may result in confusion for participants especially for dyslexic participants. They also might missed the menu items in the vertical left side of the web site as previous research (McCarthy et al., 2004), which most participants glance in the middle or top of the web page during navigation in their first visit to find target link.

As both organisations of main menus are commonly used in the web, web-experienced participants probably had encountered both organisations therefore it is not surprising that the present study failed to find significant results on the navigation performance measures. As their performance was averaged across all tasks, there is also a possibility that participants might have had different eye gaze behaviour in the first task as they became familiar with the menu structure and they later adapted themselves to the menu for other tasks. In a study with position of menus for people with learning disabilities, menus placed either horizontally on top of the page or vertically on the left of the page had no effects on time taken to find a menu item (Williams & Hennig, 2015a). In addition, findings from McCarthy et al. (2004) on non-disabled users shown that participants had different eye gaze behaviour in their first visit to a web site than their second visit to the same web site.

For participants' opinions on the organisation of menus they experienced, Unified menus was given significantly better (i.e. lower) ratings on ease of use, ease of remembering and ease of learning than Fragmented menus. Unified menus were probably easier because all of the menu items were organised in one place. Not only that, when participants were asked to rate the ease of use, ease of remembering and ease of learning for all the menus combinations used in the study, Unified menus were rated easier.

5.3.4.2 Effects of Visibility of Sub-Menus

In a study on menus with non-disabled users, Aaltonen et al. (1998) found that participants commonly read (authors used the term 'sweep', referring to the sequence of saccades that move in the same direction) items in the dynamic sub-menu (presumably top-down direction), then make quick bottom-up sweeps when searching. The present study found that menus had lower numbers of revisits when combined with Dynamic sub-menus compared to Visible sub-menus. In this study, revisit refers to a total glance returned to the particular AOI from any area outside the AOI. One possible explanation for this result is that participants read the main menu horizontally before hovering their mouse to look inside the sub-menus. While this present study did not record mouse cursor movement in parallel to participants' eye gaze, this study speculates the relation of total revisit and mouse cursor based on a study from Aaltonen et al. (1998). After reading the dynamic sub-menus in top-down direction, participants probably did not glance back (revisit) the main menus. Without giving attention to the cursor, they may move the cursor from one menu item to another item in the main menu in order to read the sub-menus. Since all of the participants were experienced Web users, they would have no problems to move the cursor on an approximate straight line following the reading direction, without the need to actually fixate at the mouse cursor in the main menus. For menu with Visible sub-menus, participants read the menu horizontally one by one and if there are sub-menus, they read vertically in top-down direction. Then, they probably return (or revisit) their eye gaze to the main menu again to continue reading the next items in the main menu.

Figure 5.14 illustrates this possible pattern of eye gaze behaviour in the Unified-Visible menu condition. From the figure, the dotted rounded-rectangles indicates the main menu items, the numbered squares indicate the sequence of eye gaze behaviour, the blue arrows indicate eye gaze movement of reading and the green arrows indicate revisits to the main menus.

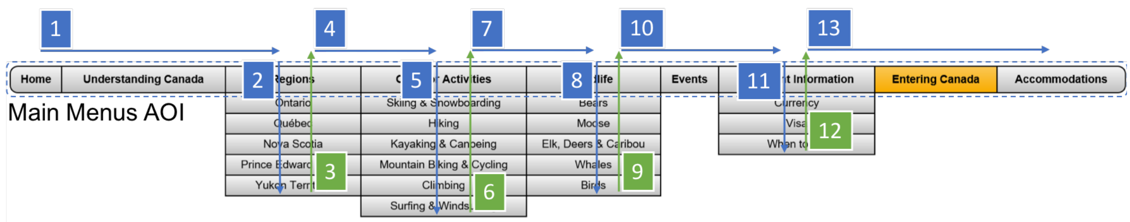


Figure 5.14 Possible revisit pattern in the Unified-Visible menu condition

Figure 5.15 illustrates the possible pattern of eye gaze behaviour in Unified-Dynamic menu condition which has a similar visual description as in Figure 5.14. By using Unified main menus as the examples, it can be seen from both Figures 5.14 and 5.15 that main menus with Visible sub-menus requires a higher number of revisits than main menus with Dynamic sub-menus.

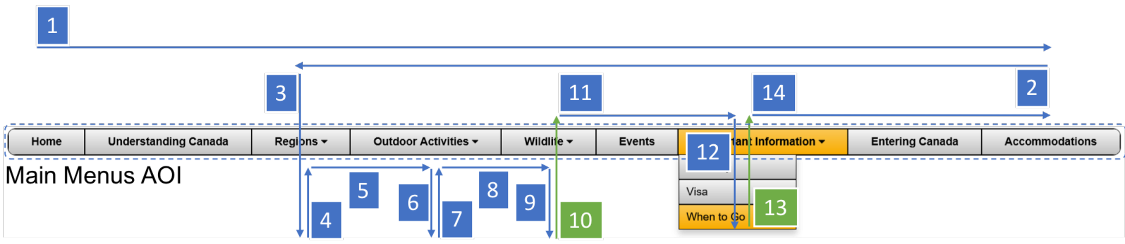


Figure 5.15 Possible revisit pattern in the Unified-Dynamic menu condition

The present study also found that participants noticed the sub-menus faster when they were Visible sub-menus compared to Dynamic sub-menus. One possible explanation for this result is participants were able to read all menu items in Visible sub-menus without the need to hover with mouse thus enabling participants to access the menu items faster. In contrast with Dynamic sub-menus, participants probably read items on the main menus first before hovering their mouse to look at the sub-menus.

5.3.4.3 Overall Effects of Participant Group

This present study found that both non-dyslexic and dyslexic participants were able to find the information they needed to answer the questions on the website (the median scores for both groups indicated that typically they answered all the questions correctly). However, dyslexic participants required longer time to complete the tasks than non-dyslexic participants. Furthermore, participants with dyslexia required more time than non-dyslexic participants in deciding whether to click on any target link that they thought would contain the information they needed. A similar result was found by Al-Wabil et al. (2008) in which reaction times of dyslexic participants were slower than non-dyslexic participants. In addition, this study found that, dyslexic participants dwelled longer and fixated more on the sub-menus compared to non-dyslexic participants. As this present study found dyslexic participants dwelled longer and made more fixations in the sub-menus, that is probably best explanation on why dyslexic participants took longer time to make their first mouse click on any target link.

According to Al-Wabil et al. (2008), dyslexic participants revisited target links three times before actually clicking on the link, while non-dyslexic participants did not revisit the target link. However, the study had small number of and no statistical test was used. It showed that dyslexic participants hesitated to click on the target link. In this present study, no significant effect on the number of revisits was found between both participant groups. However, a marginal trend (p = .05) was found whereby dyslexic participants had a higher number of revisits on the target link in the Fragmented-Visible condition compared to non-dyslexic participants. As the main menu were separated into two different areas and all items in sub-menus were visible all the time, dyslexic participants probably had to make more effort and were confused. Dyslexic participants may felt hesitant to click on the target link and need to reconfirm with all other menu items in the menu, so they probably do that and revisit the target link. As this result was marginally significant, further studies with bigger sample sizes and a wider variety of websites and tasks should be conducted with this type of menus.

Participants were asked to rate the menu combination which they experienced twice. On the first occasion, they were asked to rate the menu on ease of use, ease of remembering and ease of learning. Later, they were asked to rate again the same menu combination along with the other three menu combinations, after a chance to browse the experimental web site using different types of menu combinations. This study found no significant differences on ratings given to their assigned menus. This shows their initial opinions given on the menus was consistent.

5.3.4.4 Limitations of This Study

Limitation of this study is all four experimental tasks were completed with one web site. There is a possibility of repetition effect, in which participants may have memorised the menu combination used in the web site. However, I tried to reduce the repetition effect by averaging all the tasks into one.

Similar limitations as mentioned in previous studies, majority of participants in this study was university student. In addition, all of them had an experience of using Web more than six years. They probably encountered variety of organisation and visibility of menus on web sites.

In this study, group for dyslexic participants is different from previous studies on text presentation (discussed in Chapter 3 and Chapter 4). In previous studies, dyslexic participants were categorised into mild and moderate dyslexia, depending on the severity. With between-participants design, splitting dyslexic participants into several groups will require additional participants with at least 32 participants more.

Another limitation of participants in this study had a relatively small number of participants in each condition. Even though the total number of participants was 64, due to the need to use a between-participants design, each condition had only 8 dyslexic and 8 non-dyslexic participants. With this limited number of participants per condition, the tendency to have non-normal distribution of data is high (Krithikadatta, 2014). Between-participants design was used to avoid carry-over effect if they repeat the same tasks with other menu combinations in repeated measures design.

In addition to the limitation of the relatively small number of participants, a majority of participants in this study were recruited from my previous studies on text presentation. However, participants in this study undertook very different tasks from the previous studies, which involve navigating a website and finding information in comparison to reading texts. Therefore, their previous participation would not have influenced the results of this study.

In this study, I did not record the participants' mouse movements during the tasks. With the mouse movements recorded in addition to eye gaze data, further effects of menu organization and visibility could have been explored, such as number of revisits to the main menus. The mouse movement log could also validate the eye gaze behaviour and mouse cursor movement discussed in Figure 5.14 and Figure 5.15. In addition, initial mouse cursor position when participants were directed to the experimental web site also probably influenced the results of First Mouse Click Times variable. On occasion where some participants appeared to have memorise the position of the menu items in the menu, therefore they managed to relocate the mouse cursor to an area nearer to the target link while waiting for SMI Experiment Center™ to execute the browser

In addition, the menus in this study were designed to fit into one screen so that participants did not have to scroll around the webpages. Participants might have different results with the Fragmented main menus if it were longer than screen length as it can increase their effort and confusion by scrolling up and down in the web site. Not only that, there might be an increment in the number of revisits when the longer Fragmented main menus were combined with Visible sub-menus.

5.3.4.5 Conclusions

This study has shown that unified organisation of menus on websites produced better ratings on ease of use, ease of remembering and ease of learning. In addition, unified organisation makes participants faster in finding the appropriate target link to find information. However, web designers need to be cautious when using unified menu with visible sub-menus as participants had a higher number of fixations. If there is a need to display a large number of menu items at once, fragmented organisation with visible sub-menus is a better option to display all menu items.

With the very limited number of previous studies found which investigated the effects of the organisation and visibility of menus for people with dyslexia, this present study can be the basis for future research.

CHAPTER 6: OVERALL DISCUSSION AND CONCLUSIONS

6.1 Introduction

This chapter summarises the overall conclusion of each study conducted as part of this programme of research and outlines future lines of research. In addition, this chapter also discusses the implications and contributions of each study.

6.2 Overview of the Programme of Research

The number of people with a disability is continuing to increase due to variety of reasons. Dyslexia is a specific learning difficulty which is categorised as a cognitive disability. Approximately 5% - 10% of people in the world have dyslexia (Vidyasagar & Pammer, 2010; Washburn et al., 2014). According to WebAIM (2016), cognitive disability is one of five major categories of disabilities which affect people's use of the World Wide Web and digital technologies. The Web has grown since the early 1990s since it invention by Tim Berners-Lee (Andrews, 2018). Today, around 46% (around 3.5 billion) of the world population have access to the Web (Internet Live Statistics, 2018). As people with dyslexia have problems mainly with reading, it is crucial to present web resources in a manner that is optimised for them. In addition, such optimisation may also benefit non-dyslexic readers (Zarach, 2002). With the growth of the Web, not only is presentation of Web resources crucial but how to navigate to Web resources is also important. According to Al-Wabil et al. (2007), highly textual content and poor navigation structures can create problems for people with dyslexia.

This programme of research aimed to investigate to what extent text presentation and web menu organisation affects adults with dyslexia in comparison to adults without dyslexia in terms of their eye gaze behaviour, reading performance, as well as their subjective preferences and opinions.

For the text presentation aspect of the main research question, many variables could have been considered. However, to investigate all of the textual variables simultaneously is complicated and there were also constraints that needed to be considered. Therefore, the particular aspects chosen to investigate were typeface, font size, line spacing and line length. These were selected based on the discussion presented in Chapter 2 in which I found most of the guidelines for people with dyslexia were derived from a very small corpus of research.

Most of the definitions of dyslexia characterise reading as one of the main difficulties people with dyslexia have (Lyon et al., 2003; Reid & Green, 1996; Rose, 2009; Thomson & Watkins, 1990). Special typefaces have been designed to help people with dyslexia read more easily, including Lexie Readable, OpenDyslexic and Dyslexie (BDA Technology, 2016). As discussed in Chapter 2, mixed results have been found on the effects of typefaces on reading with people with dyslexia and some studies shown that dyslexia-optimised typefaces produced no better performance in reading than standard typefaces (Harley et al., 2013; Rello & Baeza-Yates, 2013, 2016; Wery & Diliberto, 2017). For font size, most of the published studies about people with dyslexia used absolute values of font size such as point size (Rello & Baeza-Yates, 2015; Rello et al., 2012b; Rello, Pielot, et al., 2013). One of the challenges in using absolute point size in text presentation is that size is not scalable and is rendered inconsistently across different platforms. In addition, different typefaces have different heights when used with the same point size (Aderinokun, 2015).

For line spacing, small increments in spacing can have an impact on reading performance for poor readers and people with dyslexia (Levi, 2008; Martelli et al., 2009). Most of the research on spacing for people with dyslexia has focused on character spacing and word spacing, while little attention has been given to line spacing. Most of the guidelines on text presentation for dyslexic readers recommend line spacing of 1.5 to 2.0, however results from Rello & Baeza-Yates (2015) shown that line spacing had no effect on reading by people with dyslexia. In addition, mixed results were found by other studies (Rello et al., 2012b; Rello, Pielot, et al., 2013; Schneps, Thomson, Sonnert, et al., 2013; Schoonewelle, 2013).

For line length, I found a limited number of studies on line length for people with dyslexia (Rello & Baeza-Yates, 2015; Rello et al., 2012b; Schneps, Thomson, Sonnert, et al., 2013). It is important to identify an appropriate line length for reading, since it will make the return sweep to the beginning of the next line harder if it is too long (Dyson, 2004; Dyson & Haselgrove, 2001). Studies on line length with dyslexic participants reading from a computer screen have recommended different length where shorter length (Schneps, Thomson, Sonnert, et al., 2013) and longer length (Rello et al., 2012b) than the recommendations given in the guidelines for text presentation for people with dyslexia. Most of the guidelines recommend using line lengths between 60 to 70cpl (British Dyslexia Association, 2018b; de Santana et al., 2012; Evett & Brown, 2005).

Based on the research discussed on text presentation, it can be seen that there was no conclusive evidence on what text presentations can help people with dyslexia with reading. In addition, a majority of the published studies were conducted with languages other than English (de Leeuw, 2010; Kuster et al., 2017; Pijpker, 2013; Rello & Baeza-Yates, 2013, 2015, 2016; Rello et al., 2012b; Schoonewelle, 2013; Zorzi et al., 2012), while studies with English speaking dyslexic participants were conducted with children and high school students (O'Brien et al., 2005; Schneps, Thomson, Sonnert, et al., 2013; Wery & Diliberto, 2017), not adults.

For the web navigation aspect of the main research question, I found very limited published research on people with disabilities, particularly dyslexia. With limited previous research on web navigation, I decided to investigate the effect of web menus in terms of their organisation and visibility. In the Web, menus are an important element that helps users to locate and retrieve information. The inability to locate information can frustrate people with dyslexia (Al-Wabil et al., 2007). Therefore, well-designed navigation can help in reducing their frustration. By investigating menus for people with dyslexia, this programme of research will provide a good foundation for other researchers on this topic in the future.

According to literature review on web navigation presented in Chapter 2 (see Section 2.4), people with dyslexia need a consistent structure of menus on web sites. In addition, research has shown they prefer visible menus because they had negative experiences with dynamic menus as it was difficult for them to control cursor movements when hovering over the menus (Al-Wabil et al., 2007). Guidelines for people with dyslexia recommend simple, consistent and visible navigation menus, label links clearly and to avoid dynamic menus or menus that use transparency, where dyslexic found it hard to differentiate menu items and partially visible background (de Santana et al., 2012; Evett & Brown, 2005; Friedman & Bryen, 2007; Zarach, 2002). Not only for web users with dyslexia, dynamic menus are recommended for all web users (Nielsen, 1999). On the other hand, Fowler and Stanwick (2004) recommend displaying as many menu items as possible in logical groups on a single screen. Friedman & Bryen (2007) recommend customisable placement of menus for people with dyslexia but evidence from some studies with non-disabled users has shown that users have better performance with menus placed horizontally at the top of the web page (Murano & Oenga, 2012; Murano & Sander, 2016; Williams & Hennig, 2015b).

Another aspect of interest in this programme of research is whether there are any differences between adults with different severity levels of dyslexia. In order to obtain detailed, quantifiable data as a measurement, this programme of research will use an eye tracking device to get insight of participants' behaviour and performance.

Based on the factors summarised above, three studies were formulated. Two studies were planned to investigate the text presentation variables of typeface, font size, line spacing and line length. One study was planned to investigate several variables of web menu presentation, particularly menu organisation and visibility. The methods for each study are discussed below.

Study 1: Effects of typeface and font size on detailed reading

This study was described in detail in Chapter 3. This study was designed to answer Research Questions 1(a) and 1(b).

1 (a) To what extent does the typeface used in the presentation of text on a computer screen affect eye gaze behaviour, reading performance, preferences and opinions of adults with dyslexia compared to adults without dyslexia?

1 (b) To what extent does the font size used in the presentation of text on a computer screen affect eye gaze behaviour, reading performance, preferences and opinions of adults with dyslexia compared to adults without dyslexia?

This study was a mixed-participants design with participant group as the between participants variable and typeface, typeface examples and font size as the repeated measures variables. Participants were divided into three groups on the basis of a diagnostic checklist for dyslexia (Snowling et al., 2012). The dependent variables were grouped into three types of measurements: (1) eye gaze behaviour, (2) reading performance and (3) participants' preferences and opinions. Details on the three participant groups, three levels in the typeface variable, two examples of each typeface, two levels of font size variable and three groups of dependent variables are described in Table 6.1.

Total of 36 participants took part in this study, 12 participants in each group (non-dyslexic, mild dyslexic and moderate dyslexic). All participants were native English speakers with normal vision or vision correctable with eyewear. Participants read 12 experimental texts, one with each combination of typeface example and font size. After reading all the experimental texts, the participant completed the post-study questionnaire.

Study 2: Effects of line spacing and line length on detailed reading

This study was described in detail in Chapter 4. This study was designed to answer Research Questions 1(c) and 1(d).

1 (c) To what extent does the line spacing used in the presentation of text on a computer screen affect eye gaze behaviour, reading performance and preferences of adults with dyslexia compared to adults without dyslexia?

1 (d) To what extent does the line length used in the presentation of text on a computer screen affect eye gaze behaviour, reading performance and preferences of adults with dyslexia compared to adults without dyslexia?

This study was a mixed-participants design with participant group as the between participants variable and line spacing and line length as the repeated measures variables. Participants were divided into three groups on the basis of a diagnostic checklist for dyslexia (Snowling et al., 2012). The dependent variables were grouped into three types of measurements: (1) eye gaze behaviour, (2) reading performance and (3) participants' preferences. Details on the three participant groups, two levels in the line spacing variable, two levels of line length variable and three groups of dependent variables are described in Table 6.1.

Total of 24 participants took part in this study, 8 participants in each group (non-dyslexic, mild dyslexic and moderate dyslexic). All participants were native English speakers with normal vision or vision correctable with eyewear. Participants read four experimental texts, one with each combination of line spacing and line length. After reading all the experimental texts, the participant completed the post-study questionnaire.

Study 3: Effects of menu organisation and visibility on web navigation

This study was described in detail in Chapter 5. This study was designed to answer Research Questions 2(a) and 2(b).

2 (a) To what extent does menu organisation in web navigation affect eye gaze behaviour, navigation performance and opinions of adults with dyslexia compared to adults to without dyslexia?

2 (b) To what extent does menu visibility in web navigation affect eye gaze behaviour, navigation performance and opinions of adults with dyslexia compared to adults to without dyslexia?

This study was a between-participants design with participant group, organisation of navigation menus and visibility of navigation menus as the between participants variables. Participants were divided into two groups on the basis of a diagnostic checklist for dyslexia (Snowling et al., 2012). The dependent variables were grouped into three types of measurements: (1) eye gaze behaviour, (2) navigation performance and (3) participants' opinions. Details on the two participant groups, two levels in the menu organisation variable, two levels of menu visibility variable and three groups of dependent variables are described in Table 6.1.

Total of 64 participants took part in this study, 32 participants in each group (non-dyslexic and dyslexic). All participants were native English speakers with normal vision or vision correctable with eyewear. Participants complete four experimental tasks in which they navigated in a web site to find answer for the question asked in each task. They were assigned one menu condition with a combination of menu organisation and visibility. After completing all the experimental tasks, the participant completed two post-study questionnaires. In the first questionnaire, they rated the menu combination they had experience. In the second questionnaire, they rated all four menus used in this study. Both questionnaires asked for ratings of ease of use, ease of remembering and ease of learning.

Table 6.1 Summary of methods for the three studies in this programme of research

	TEXT PRESENTATION		**WEB NAVIGATION**
	STUDY 1	**STUDY 2**	**STUDY 3**
Design	3 × 3 × 2 × 2 mixed-participants	3 × 2 × 2 mixed-participants	2 × 2 × 2 between-participants
Participant Groups	Non-dyslexic Mild dyslexic Moderate dyslexic	Non-dyslexic Mild dyslexic Moderate dyslexic	Non-dyslexic Dyslexic

Independent Variables	Typeface and its two examples: 1. Serif: a. Times New Roman b. Georgia 2. Sans Serif: a. Verdana b. Arial 3. Dyslexia-optimised: a. Lexie Readable b. OpenDyslexic Font size: 1. Small (2.5mm) 2. Medium (3.3mm)	Line spacing: 1. Spacing of 1.5 (150% of 18-point size, 27-point size) 2. Spacing of 2.0 (200% of 18-point size, 36-point size) Line length: 1. 60 to 70cpl 2. 80 to 90cpl	Menu organisation: 1. Unified (menus are located on top of the page) 2. Fragmented (menus are located on top and left side of the page) Menu visibility: 1. Visible (sub-menus are static and visible) 2. Dynamic (sub-menus appear when hovered with mouse)
Dependent Variables	Eye gaze behaviour: 1. Number of fixations 2. Average fixation duration	Eye gaze behaviour: 1. Number of fixations 2. Average fixation duration	Eye gaze behaviour: 1. Entry time 2. Dwell time 3. First fixation duration 4. Number of revisits 5. Number of fixations 6. Average fixation duration
	Reading performance: 1. Comprehension score 2. Reading time	Reading performance: 1. Comprehension score 2. Reading time	Navigation performance: 1. First mouse click time 2. Task completion time 3. Navigation score
	Preferences and opinions: 1. Ratings on ease of read 2. Preferred combination 3. Ratings on interest 4. Ratings on comfort	Preferences: 1. Ratings of ease of read 2. Preferred combination	Opinions: 1. Ratings on ease of use 2. Ratings on ease of remembering 3. Ratings on ease of learning

6.3 Contributions of the Programme of Research

6.3.1 Improvements to the Methodology of Studies in Investigating Digital Text Presentation and Navigation

A first overall contribution of this programme of research is to the methodology of studies to investigate text presentation on screen and web navigation effects for people with dyslexia. As discussed in Chapter 2, some of the studies on people with dyslexia had a number of methodological problems such as small numbers of participants (Al-Wabil et al., 2007, 2008; Wery & Diliberto, 2017), participants' age range were too wide (Rello & Baeza-Yates, 2013, 2015, 2016; Rello et al., 2012b; Rello, Pielot, et al., 2013), not providing statistical results (Al-Wabil et al., 2007, 2008; de Leeuw, 2010; Rello et al., 2012b), unclear and confusing analyses (Rello & Baeza-Yates, 2013, 2016; Zorzi et al., 2012), repetition effects which participants did the same task multiple times on different session (Kuster et al., 2017), texts read by participants being too short or of inconsistent length (Rello & Baeza-Yates, 2015; Rello et al., 2012b; Rello, Pielot, et al., 2013), using a combination of different types of passages (e.g. prose and poetry) (Rello et al., 2012b), testing variables separately (Rello & Baeza-Yates, 2015; Rello et al., 2012b), using only one dependent variables to measure eye gaze behaviour (Rello & Baeza-Yates, 2012, 2015; Rello et al., 2012b; Rello, Pielot, et al., 2013). Most of the studies asked comprehension questions as part of the experimental tasks, but none of the studies measured the difficulty of the comprehension questions or whether the participants were likely to know the answers from general knowledge. If the questions are too difficult, it will affect participants' performance particularly with their scores and completion time. If the questions are too easy, they will not be a good measure of comprehension.

For the two text presentation studies in this programme of research, the themes for the texts were selected to suit university students. All of the texts were carefully constructed to have similar characteristics (e.g. number of words, paragraphs) and similar levels of difficulty. A number of readability measures (FKGL, FKRE, and GFS) were used to ensure that all the texts were comparable and at a suitable readability level for university students. Not only the texts, but also the comprehension questions asked in all the studies were carefully prepared in a manner so that the questions had similar levels of difficulty. Two different aspects of difficulty were measured, the overall difficulty of the question and difficulty of finding the answer in the text. In addition, participants' familiarity on the questions were also measured. This is to make sure that participants would answer the questions due to their understanding of the text read during the experimental session, not on their previous general knowledge on the subject matter.

People with dyslexia show varying symptoms from individual to individual with different levels of severity and different manifestations of dyslexia. In addition to this, they may show inconsistent results in different dyslexia assessments (Rello et al., 2012b; Riddick et al., 1997). Therefore, it is important to distinguish people with dyslexia into some finer grained categories. In this programme of research, I divided participants in those whose have mild and moderate dyslexia with a well-established screening tool, the Adult Reading Questionnaire (Snowling et al., 2012). This allowed me to investigate whether there are any differences due to severity of dyslexia.

For the design of the studies on text presentation, a repeated measures design was selected because it was appropriate for all the participants to experience all of the treatment conditions, as a number of different texts could be created which were matched for characteristics and level of difficulty. Repeated measures is a powerful design since factors that can cause variability between participants are controlled for. In addition, fewer participants are needed to achieve acceptable effect sizes as all the participants are involved with multiple treatment conditions. The order of the texts was counterbalanced to reduce practice and fatigue effects. In Study 2, not only the sequence of texts was counterbalanced, but each text was counterbalanced across all line spacing and line length condition.

For Study 1, two examples of typefaces were used to represent each of the serif, sans serif and dyslexia-optimised families of typefaces. All studies investigating the effects of typefaces on people with dyslexia (de Leeuw, 2010; Kuster et al., 2017; Pijpker, 2013; Rello & Baeza-Yates, 2013, 2016; Wery & Diliberto, 2017) have used only one dyslexia-optimised typeface. With two examples of typefaces, the possibility of results influenced by specific typeface was reduced. For font size, this study used physical height on screen instead of using absolute size (such as point size) as most of the font size studies for people with dyslexia have done (Rello & Baeza-Yates, 2015; Rello et al., 2012b; Rello, Pielot, et al., 2013). This was to avoid the results being contaminated with variety of heights on same point size for different typefaces. A character with Arial 14-point size (representing Small font size) and 18-point size (representing Medium font size) were printed and carefully measured as a baseline.

For the web navigation study in this programme of research, materials were prepared by choosing a highly reputable source, Lonely Planet (2017) which has higher customer ratings, 4.9/5.0 stars (in Amazon.co.uk), 3.7/5.0 stars (in Amazon.com) and 4.0/5.0 stars (in goodread.com) as a reference. Based on the source, the logical structure of the web site and its content were developed. The web site and user tasks were also tested in a preparatory study in order to measure the difficulty of finding answers for the questions. In terms of locating the answer in the web site, all of the questions have similar numbers between easy and difficult questions. A balanced number of easy and difficult questions would help even out the effort of participants. In addition, participants' potential familiarity with the questions was also measured. Limited studies were found on web navigation with people with dyslexia (Al-Wabil et al., 2007; MacFarlane et al., 2010). Therefore, organisation and visibility of menus were chosen to be investigated as menus are an important element in doing navigation.

For the design of the study on web navigation, a between-participants design was selected. As participants need to complete four tasks with one menu combination on the same web site, a carry-over effect will be introduced if they repeat the same tasks with other menu combinations. In order to minimise the effect, between-participants design was chosen as in this case the treatment conditions were given to different groups of participants. In order to reduce the effects of differences between individual participants to some extent, participants' age and gender were controlled in each treatment condition.

For participants in the web navigation study, only two groups were considered, which is adults without dyslexia and with dyslexia. Based on the results of text presentation studies, no significant differences (see Table 6.3) were found between participants with mild and moderate dyslexia on different typefaces, font sizes, line spacings and line lengths. In addition, no consistent significant differences were found between participants with mild and moderate dyslexia. A summary of this results can be seen in Table 6.2. Participants with moderate dyslexia had more fixations in Study 1, longer fixation durations and reading times in Study 2 compared to participants with mild dyslexia. Since no major differences were found between the two dyslexic groups, I decided to combine them as one group in Study 3, as having two separate dyslexic groups would require additional numbers of participants. For participants with dyslexia, each treatment condition had almost similar number of participants with mild and moderate dyslexia. Even though participants with mild and moderate dyslexia were grouped together, their severity were considered in order to have a representative sample of dyslexic participants .

6.3.2 Demonstration of the Variability of Results

For these carefully designed studies investigating a range of text presentation on screen and web navigation effects for people with dyslexia, findings are summarised in Table 6.2 – 6.4. Table 6.2 shows the findings on the overall effects of participant group, while Table 6.3 shows the findings from the objective measures and Table 6.4 shows the findings from the subjective measures.

Table 6.2 Summary of significant group level differences between non-dyslexic and dyslexic participants for eye gaze behaviour and performance

1. Text presentation			
a.	Eye gaze behaviour	**Number of fixations**	
		Study 1:	Overall: Moderate dyslexic > Non-dyslexic
		Study 2:	Overall: No differences
		Fixation duration	
		Study 1:	Overall: No differences
		Study 2 :	Overall: Moderate dyslexic > Mild dyslexic
b.	Reading performance	**Reading time**	
		Study 1:	Overall: No differences
		Study 2:	Overall: Moderate dyslexic > Non-dyslexic / Mild dyslexic
		Comprehension score	
		Study 1:	Overall: No differences
		Study 2:	Overall: Mild dyslexic/Moderate dyslexic > Non-dyslexic
2. Web navigation			
a.	Eye gaze behaviour	**Number of fixations**	
		Study 3:	Overall: For sub menus, dyslexic > Non-dyslexic
		Dwell time	
		Study 3:	Overall: For sub menus, dyslexic > Non-dyslexic
b.	Navigation performance	**First mouse click time**	
		Study 3:	Overall: Dyslexic > Non-dyslexic
		Task completion time	
		Study 3:	Overall: Dyslexic > Non-dyslexic

Note. Study 1 – the effects of typeface and font size on detailed reading. Study 2 – the effects of line spacing and line length on detailed reading. Study 3 – the effects of menu organisation and menu visibility on web navigation.

Table 6.3 Summary of significant differences due to text and menu presentation manipulations of non-dyslexic and dyslexic participants

1. Text presentation		
a.	Typeface	**Serif typeface** • Had fewer fixations with small font size than medium font size • Small serif typefaces had fewer fixations than small dyslexia-optimised typefaces **Sans serif typeface** • Had fewer fixations with small font size than medium font size **Dyslexia-optimised typeface** • Had fewer fixations with small size than medium size • Dyslexia-optimised typefaces had shorter fixations than serif and sans serif typefaces • Non-dyslexic and moderate dyslexic participants had shorter fixations with dyslexia-optimised typefaces than serif and sans serif typefaces
b.	Font size	**Small size** • Had fewer fixations than medium size • All participants had fewer fixations with small size than medium size **Medium size** • Had shorter fixation than small size • Had fewer fixations with non-dyslexic than moderate dyslexic participants
c.	Line spacing	**1.5 versus 2.0** • No differences
d.	Line length	**60-70cpl versus 80-90cpl** • 80-90cpl had fewer fixations than 60-70cpl • 80-90cpl had shorter reading time than 60-70cpl
2. Web navigation		
a.	Organisation	**Unified versus Fragmented** • Target link had shorter entry time in unified organisation than fragmented organisation • Main menus with visible sub-menus had fewer fixations in fragmented organisation than unified organisation
b.	Visibility	**Visible** • Visible sub-menus had shorter entry time than dynamic sub-menus **Dynamic** • Main menus had fewer revisits with dynamic sub-menus than visible sub-menus • Unified main menus had fewer fixations with dynamic sub-menus than visible sub-menus

Table 6.4 Summary of significant differences for preferences and opinions of non-dyslexic and dyslexic participants

1. Text presentation		
a.	Typeface	**Sans serif typeface** • Preferred by 50% of all participants • Rated as moderately easy to read • Non-dyslexic and mild dyslexic participants rated sans serif as easier to read than dyslexia-optimised typefaces **Serif typeface** • Preferred by 36% of all participants • Rated as moderately easy to read • Non-dyslexic and mild dyslexic participants rated serif as easier to read than dyslexia-optimised typefaces **Dyslexia-optimised typeface** • Preferred by 14% off all participants • Rated as slightly difficult to read • Non-dyslexic participants felt neutral about receiving documents in Lexie Readable or OpenDyslexic • Dyslexic participants felt uncomfortable in submitting documents in OpenDyslexic • Dyslexic participants rated OpenDyslexic as moderately difficult to read compared to standard typefaces • Dyslexic participants felt slightly uninterested in using OpenDyslexic
b.	Font size	**Small** • Moderate dyslexic participants rated small size as slightly difficult to read **Medium** • Moderate dyslexic participants rated medium size as slightly easy to read
c.	Line spacing	**1.5** • Rated as slightly easy to read **2.0** • Rated as moderately easy to read • Non-dyslexic participants with 80-90cpl rated line spacing 2.0 as easier to read than line spacing 1.5
d.	Line length	**60-70cpl versus 80-90cpl** • No difference on ease of reading
2. Web navigation		
a.	Organisation	**Unified** • Rated moderately easy to use, to remember and to learn **Fragmented** • Rated as slightly easy to use, to remember and to learn
b.	Visibility	**Visible versus Dynamic** • No difference on ease of use, remembering and learning

Study 1 and Study 2: Effects of text presentation (typeface, font size, line spacing and line length) on detailed reading

There is good evidence that dyslexic participants show poorer eye gaze behaviour in reading, with more fixations and longer fixation durations (Rayner, 1998). From Table 6.2, it can be seen that studies with text presentation adding more evidence in that participants with dyslexia had poorer performance than participants without dyslexia in terms of number of fixations and fixation durations. From Table 6.3 and Table 6.4, it can be seen that typeface, font size, line spacing and line length had no major significant effects on the three participant groups. As claimed by Rayner et al. (2012), there was no eye gaze differences in reading when given reasonable typeface, font size, line length and line spacing.

Findings on objective measures from Study 1 (typeface and font size) showed that small font size (approximately similar height with Arial 14-point size) had fewer fixations and longer fixation duration in comparisons to medium font size (approximately similar height with Arial 18-point size) for all participants. In addition, dyslexia-optimised typefaces had the shortest fixation duration compared to serif and sans serif. Subjective measures showed that sans serif typefaces are preferred by all participants and moderate dyslexic participants rated medium font size as easier to read. Participants with dyslexia were not interested or comfortable using dyslexia-optimised typefaces while participants without dyslexia were neutral in receiving documents using these typefaces.

Findings from Study 2 (line spacing and line length) showed that objective measures support longer line length of 80-90cpl as better for all participants, while subjective measures support wider line spacing of 2.0 as easier for all participants.

Study 3: Effects of menu organisation and visibility on web navigation

There are very few studies published on web navigation for people with dyslexia (Williams & Hennig, 2015a). This study contributes by providing a foundation for future research on this topic and also gives useful insights on menu organisation and visibility. From Table 6.2, it can be seen that findings from this study showed that participants with dyslexia react more slowly than participants without dyslexia in terms of making their first mouse click in the web site. In addition, participants with dyslexia completed the task more slowly, dwelt longer on menu items and had more fixations.

Findings from the objective measures showed that participants noticed the target link to the correct web page on which to find the answer to the question faster with a unified organisation while fragmented organisation with visible sub-menus and unified organisation with dynamic sub-menus had fewer fixations. In addition, dynamic sub-menus had fewer revisits. On the subjective measures, participants rated the unified menu organisation as easier to use, to remember and to learn than the fragmented menu organisation.

6.3.3 Contributions to the Development and Updating of Web Accessibility Guidelines for People with Dyslexia

Findings from these studies on text presentation and web navigation will contribute to the development and updating of guidelines about on-screen text presentation and website organisation for people with dyslexia. Table 6.5 shows the recommendations based on the results of objective and subjective measures in the studies of text presentation and web navigation. Despite of these recommendations, people with dyslexia need to be given flexibility in choosing which specific typeface, font size, line spacing and line length that may benefit their reading. In addition, recommendations made for navigation are based on a study with limited previous literature and limited number of participants. As discussed previously, this study clearly needs more future work before proposing more firm recommendations.

Table 6.5 Recommendations made for people with dyslexia based on the text presentation and web navigation studies

Text Presentation	
Categories	**Recommendation**
Typeface	Typefaces in Sans Serif and Serif
Font Size	Small size (approximately similar height with Arial, 14pt font) and above
Line Spacing	1.5 and above
Line Length	80 – 90cpl
Web Navigation	
Categories	**Recommendation**
Menu Organisation	Unified
Menu Visibility	Depending on the menu organisation • Use dynamic menus with unified organisation • Use visible menus with fragmented organisation

6.3.4 Other Contributions

In addition to the above contributions, this programme of research contributes to the research on digital text presentation and providing good foundation for future research in web navigation for people with dyslexia. Furthermore, this programme of research provides empirical insight that university students with mild and moderate dyslexia had few objective differences in their reading performance in relation to typeface, font size, line length and line spacing in comparison to university students without dyslexia.

6.4 Discussion of Research Questions

Having discussed the findings for research questions 1(a) – 1(d) and 2(a) – 2(b) of each study in Section 3.3.4.1, Section 3.3.4.2, Section 4.3.4.1, Section 4.3.4.2, Section 5.3.4.1, and Section 5.3.4.2; this section will discuss additional research questions 3(a) – 3(d) based on severity levels of dyslexia made in Section 1.8.

3 (a) To what extent will adults with mild dyslexia have more efficient eye gaze behaviour, better reading performance and better ease of read on dyslexia-optimised typefaces, in comparison to adults with moderate dyslexia

In Study 1, significant interaction was found between participant group and typeface. Participants with mild dyslexia had no significant effect on eye gaze behaviour and reading performance when reading with dyslexia-optimised typefaces. Furthermore, they had better ease of read rating with serif and sans serif typefaces than dyslexia-optimised typefaces. In addition, participants with moderate dyslexia had shorter fixation duration with dyslexia-optimised typeface. Therefore, adults with mild dyslexia do not have more efficient eye gaze, better reading performance and better ease of read on dyslexia-optimised typefaces, in comparison to adults with moderate dyslexia.

3 (b) To what extent will adults with mild dyslexia have more efficient eye gaze behaviour, better reading performance and better ease of read on larger text, in comparison to adults with moderate dyslexia

In Study 1, significant interaction was found between participant group and font size. Both participants with mild dyslexia and moderate dyslexia had significant fewer number of fixations with smaller text. However, participants with mild dyslexia had no significant fewer number of fixations than participants with moderate dyslexia with the size. In addition, participants with moderate dyslexia had better ease of read rating on larger text. Therefore, adults with mild dyslexia do not have more efficient eye gaze, better reading performance and better ease of read on larger text, in comparison to adults with moderate dyslexia.

3 (c) To what extent will adults with mild dyslexia have more efficient eye gaze behaviour, better reading performance and better ease of read on wider line spacing, in comparison to adults with moderate dyslexia

In Study 2, no significant interaction was found between participant group and line spacing on eye gaze behaviour, reading performance and ease of read. Therefore, adults with mild dyslexia do not have more efficient eye gaze, better reading performance and better ease of read on wider line spacing, in comparison to adults with moderate dyslexia.

3 (d) To what extent will adults with mild dyslexia have more efficient eye gaze behaviour, better reading performance and better ease of read on shorter line lengths, in comparison to adults with moderate dyslexia

In Study 2, no significant interaction was found between participant group and line length on eye gaze behaviour, reading performance and ease of read. Therefore, adults with mild dyslexia do not have more efficient eye gaze, better reading performance and better ease of read on shorter line lengths, in comparison to adults with moderate dyslexia.

6.5 Limitations and Future Work

While this programme of research has provided an investigation of digital text presentation and web navigation by adults with and without dyslexia, there are a number of limitations in the programme of research which need to be discussed. The fact that a majority of participants recruited were university students and experienced Web users, means caution must be exercised when generalising findings found from all three studies in this programme. As university students, the participants with dyslexia may have developed strategies to help them read well in comparison to people with dyslexia who have not achieved well enough academically to enter higher education (Grigorovich-Barsky & Belson, 2013). In addition, both dyslexic and non-dyslexic participants were very experienced in using the Web, they would have encountered a variety of organisation and visibility of menus on web sites.

A more fundamental limitation is the nature of the eye gaze data. Readers typically process information during fixation (Ellis, 2016; Rayner et al., 2012). Evidence suggests that dyslexic readers had comparable eye gaze with non-dyslexic readers when given text appropriate to their reading ability while non-dyslexic readers had comparable eye gaze with dyslexic readers when given difficult text (Hyönä & Olson, 1995; Rayner, 1998). Present studies showed that participants with dyslexia had more fixations than participants without dyslexia in some part of the text presentation aspect, which was evidence in the study of typeface and font size (Study 1) but not in line spacing and line length study (Study 2). As discussed in Rayner (1998), poor eye gaze manifests underlying problem in reading and not a cause of reading disability. It is inconclusive why participants with dyslexia had more fixations during reading in Study 1 but not in Study 2, although texts in both studies had similar readability as measured by FKRE, FKGL and GFS. Even though texts in both studies had similar readability, both studies had different themes whereby text in Study 1 is about Olympic Games while Study 2 is about fruit, vegetable and cities. This leads to a further question whether different themes would or would not play a role in poor eye movement.

There are also a number of aspects that can be investigated further. Future work is discussed below.

6.5.1 Further Investigations on Longer Texts and Longer Vertical Menus

While the studies in this programme of research have used longer texts than most of the studies in reading with dyslexia, the texts used did fit in one screen so that participants did not have to scroll. For the text presentation studies in this programme of research which involved detailed reading, the number of words per text are between 270 to 285 words in Study 1, between 225 to 235 in Study 2. According to Wikipedia[27], average number of words per article in Wikipedia is 320 words. Therefore, using the average words from Wikipedia is a good start to investigate further on the effects of longer texts on participants with dyslexia on detailed reading. With longer texts, participants might begin to feel tired and may have different eye gaze behaviour and performance and differences between dyslexic and non-dyslexic participants might emerge.

[27] https://en.wikipedia.org/wiki/Wikipedia:Words_per_article

In addition, investigation into vertical menus with lengths longer than a single screen could be conducted to see if they worsen the web navigation experience. Longer vertical screens may increase participants' effort and confusion as they need to scroll up and down in the web site in order to find desired menu items. They might revisit and fixate more on the menus.

6.5.2 Further Investigations on Mouse Movement in Parallel with Eye Gaze

Eye tracking is a very useful tool to obtain unbiased, objective and quantifiable data as a measurement in reading and navigation studies. While the current research programme has used eye tracking in all of the studies, recording the participants' mouse movements in parallel with eye gaze data could give yet more insight into participants' behaviour. It would have been helpful to gain more understanding, particularly in the study with web navigation. While conducting the web navigation study with people with dyslexia, I realised there are different mouse movements depending on the menu combinations. However, a study needs to be properly designed and conducted in order to understand the movements, whether it differs significantly between participant groups, menu combinations or was just random behaviour. It is also interesting to see if there are relationships between the mouse cursor and eye gaze movement for people with dyslexia when doing navigation. In a study on web browsing with non-disabled users, a strong relationship between mouse cursor and eye movement was found (Chen, Anderson, & Sohn, 2001; Huang, White, & Buscher, 2012).

6.5.3 Analysing the Scan Paths in Web Navigation

Another aspect that is worth conducting further analyses is to explore participants' scan paths when doing web navigation. A scan path is a larger unit of eye movement containing a sequence on fixations and saccades at different locations. The information in these sequences can help researchers to understand and identify a general direction and strategy which participants follow in a web site. According to Eraslan et al. (2015), there are several categories of techniques in identifying the trending scan path such as similarity/dissimilarity calculation (two scan paths are compared to determine the similarity and dissimilarity), transition probability calculation (probabilities for transition of visual elements are determined), pattern detection (searching for a pattern based on number of matches) and common scan path identification (detecting common scan path to represent entire participant group).

6.5.4 Extending the Research to Different Categories of Dyslexic Participants

The studies on text presentation in this programme of research differentiated participants with dyslexia into two groups based on their severity, mild and moderate dyslexia. However, the studies did not find any significant differences between these groups in text presentation and web navigation. One of the challenges in describing dyslexia is people show varying symptoms from individual to individual (Frith, 1999). In addition, people with dyslexia have different patterns of strength and weaknesses in reading (Ellis, 2016). There is evidence that there are different distinct sub-groups of people with dyslexia (Hanley, 2017). Therefore, future research is needed with different sub-groups of people with dyslexia. A categorisation could be made based on similar reading profiles or similar patterns of symptoms. With such different categorisations of people with dyslexia, different patterns of eye gaze behaviour may emerge.

6.6 Conclusions

Studies on text presentation and web navigation have shown that manipulation of text and menu presentation have effects on both non-dyslexic and dyslexic participants. For studies on text presentation, mixed results were found for typeface, font type, line length and line spacing. Study on typeface and font size uncovered that dyslexia-optimised typeface had slight benefit on reading with both non-dyslexic and dyslexic participants. However, from the overall results on eye gaze behaviour and reading performance, the typeface did not appear to have better readability as claimed by font designers. Moreover, majority of participants preferred to read with standard typefaces; sans serif and serif. Small and medium font size had shown mixed results in objective measures however in subjective measures, moderate dyslexic participants rated bigger font size as easier to read. Study on line spacing and line length found that participants performed better with longer line length while they rated bigger line spacing as easier to read. Based on these studies on text presentation, it would be beneficial to users if web developers provide customisation capability in their web sites where the users can choose which one works best for them. However, the choice of customisation needs to be filtered down to a range of typefaces, font sizes and line spacing that have shown significant positive results on either objective or subjective measures. Furthermore, it is advised to revise the guideline for people with dyslexia, particularly with line length where longer length should be recommended than current length. As users are becoming more competent and experienced in web, they are used to read texts from screen. In addition, most of the computers now have wider screen which provides more room for texts. However, more research is needed in order to identify at which length will there be adverse effects on users when reading with longer texts.

As far as web navigation is concerned, findings from this study found that dyslexic participants navigate sub-menus slower than non-dyslexic participants. In addition, this study found that dyslexic participants had no difference from non-dyslexic participants when menus are presented with different organisation and visibility. With lack of previous studies on web navigation, it is clear that more research could be done in this aspect especially for people with dyslexia.

APPENDIX A – CONVERSION OF LOGMAR PRINT SIZES

logMAR value	x-height (mm)	Point (pt)*	Sample Texts	Example of Usage
1.0	5.82	32.0	hello	Sub-headlines, children's book
0.9	4.65	25.0	hello	Large print books
0.8	3.64	20.0	hello	Books for ages 7-8
0.7	2.91	16.0	hello	Computer type display
0.6	2.33	12.5	hello	Grade 1-3 children's book
0.5	1.82	10.0	hello	Grade 4-6 children's book
0.4	1.45	8.0	hello	Small column newspaper
0.3	1.16	6.3	hello	Telephone directory
0.2	0.92	5.0	hello	Small print Bible, footnote
0.1	0.73	4.0	hello	Mail order catalogues
0.0	0.58	3.2	hello	Medicine bottle labels
-0.1	0.47	2.5	hello	*-nil-*
-0.2	0.36	2.0	hello	*-nil-*

Note. Conversion of logMAR print size within viewing distance of 40 cm/16 inches. Numbers were rounded to simplify sequences. Rounding errors do not exceed 1.2 percent. Adapted from National Research Council et al. (2002) and Willings (2017).

APPENDIX B – SAMPLE OF TEXT, QUIZ AND DIFFICULTY ASSESSMENT FOR STUDY 1

Practice 1 – The Modern Olympic Games

The Olympic Games are a major international sporting event in which thousands of athletes participate. They are considered the world's most prestigious sports competition with more than 200 nations participating. The Games are held every four years. Until 1992, Summer and Winter Olympics were both held in the same year. Since then, the Winter Olympics have shifted to the even years between the Summer Games which helps space out the planning of the two events. The creation of the modern Games was inspired by the ancient Olympic Games, which were held for nearly 12 centuries in Olympia, Greece.

Planning for the modern Olympic Games began in 1894 when Pierre de Coubertin formed the International Olympic Committee (IOC). The idea for reviving the Olympic Games first came to Coubertin in 1889. He spent the next five years organising an international meeting of athletes. The IOC has become the governing body of the Olympic Movement and the Olympic Charter. Its headquarters were originally based in Paris before they were moved to Lausanne in Switzerland.

As the governing body, the IOC is responsible for choosing the host city for each Olympic Games; the IOC also decides the Olympic programme. This includes the sports to be contested at each Games. The host city is responsible for organising and funding a celebration of the Games consistent with the Olympic Charter.

The Games include many important symbols, such as the Olympic flag and torch; these are particularly used in the elaborate opening and closing ceremonies. The Games have grown so much that nearly every nation throughout the world is now represented. The first, second and third placed competitors receive Olympic medals in gold, silver and bronze respectively.

Please answer these questions without looking at the text (circle one of the crosses or put an X over it).

1. Approximately how many countries participate in the Olympic Games?
    a. More than 100 nations
    b. More than 150 nations
    c. More than 200 nations

    How difficult was this question?

    Very difficult *Not at all difficult*
    \+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

2. Which of the following statements describe the responsibilities of the International Olympic Committee (IOC)?
    a. Choosing the host city for each Olympic Games
    b. Organizing the opening and closing ceremonies of the Games
    c. Managing accommodation for competitors at the Games

    How difficult was this question?

    Very difficult Not at all difficult
    \+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

3. The ancient Olympic Games were held ...
    a. for nearly 12 centuries
    b. in Athens, Greece
    c. every four years

    How difficult was this question?

    Very difficult *Not at all difficult*
    \+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

4. The Summer and Winter Olympic Games ...
    a. have been held two years apart since 1894
    b. have never been held in the same year
    c. have been held two years apart since 1992

    How difficult was this question?

    Very difficult *Not at all difficult*
    \+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

Now have a look at the text, and rate how hard you think the questions were when you can consult the text. Tick the box whether you knew the answer before reading the text.

1. Approximately how many countries participate in the Olympic Games?

Very difficult *Not at all difficult*

+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text ☐ YES ☐ NO

2. Which of the following statements describes the responsibilities of the International Olympic Committee (IOC)?.

Very difficult *Not at all difficult*

+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text ☐ YES ☐ NO

3. The ancient Olympic Games were held ...

Very difficult *Not at all difficult*

+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text ☐ YES ☐ NO

4. The Summer and Winter Olympic Games ...

Very difficult *Not at all difficult*

+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text ☐ YES ☐ NO

Practice 2 – Football in Olympic Games

Football (also known as soccer) was not on the programme at the first modern Olympic Games in 1896. International football was only in its infancy at the time. However, some sources claim that an unofficial football tournament was organised during the first Games. In that tournament an Athens team lost to a team from Smyrna (now Izmir), then part of the Ottoman Empire. Smyrna went on to win against a team from Denmark.

Football tournaments were then played at the 1900, 1904 and 1906 Games, but not by national teams. Although the International Olympic Committee (IOC) considers the 1900 and 1904 tournaments to be official events, they are not recognised by the International Federation of Association Football (FIFA). Neither organisation recognises the 1906 Games, as they were a games which were to be held in Athens between the main Games, but they were only held once.

The first full tournament was played at the London Games in 1908. Since then football has become an important part of the summer Olympic Games. The competition became more important during the 1920s, although that decade also saw a very bad day for Olympic football. During the 1920 final, Czech players walked off the field in order to raise awareness of their anger at the refereeing and the militarised mood at the Games.

In the 1924 and 1928 Olympic games, the first South American teams, Uruguay and Argentina, entered the competition. Uruguay won the competition at both these games and officials became aware that the Olympic movement was hindering the ability of nations to participate on an equal footing. The Olympics only permitted amateurs to participate, so it did not represent the true strength of the international game.

Experimental 1 – The Ancient Olympic Games

The Ancient Olympic Games were closely linked to religious festivals. They were held every four years at the temple of Zeus in Olympia, Greece. Competition involved representatives of a number of city-states and kingdoms throughout Ancient Greece. It is believed that during the Games, all conflicts among the participating city-states were postponed until the Games were finished. This cessation of hostilities was known as the Olympic peace or truce.

According to historical records, the first Ancient Olympic Games can be traced back to 776 BC. They were dedicated to the Olympian gods and were staged on the ancient plains of Olympia. They continued for nearly 12 centuries, however there is no consensus as to when the Ancient Olympic Games ended. The Ancient Olympic Games gradually declined in importance as the Romans gained power and influence in Greece.

The origin of the Olympics is mysterious and myths identify Heracles and his father Zeus as the founders of the Games. According to legend, it was Heracles who first called the Games "Olympic" and he built the Olympic Stadium in honour of Zeus. Following its completion, he walked in a straight line for 200 steps and called this distance a "stadion", which later became an Ancient Greek and Roman unit of distance.

The Ancient Olympics featured sporting events alongside ritual sacrifices, which honoured Zeus and Pelops, a divine hero and mythical king of Olympia. The winners of events were admired and remembered in poems and statues. The Games featured running events, a pentathlon (consisting of a jumping event, discus throwing, foot races and wrestling), boxing, wrestling, and equestrian events. The Olympic Games were part of a cycle of games known as the Panhellenic Games.

Experiment 2 – Gymnastics at the Olympic Games

Gymnastics have been contested at every Summer Olympic Games since the birth of the modern Olympics. Initially, only men athletes were allowed to compete in this sport. However, since the 1928 Summer Olympics, women were allowed to compete, but initially only in artistic gymnastic events. Rhythmic gymnastics events were introduced at the 1984 Summer Olympics, and trampolining events were added at the 2000 Summer Olympics.

Artistic gymnastics is the best known of the events in this group with six events for male athletes: floor exercise, pommel horse, still rings, vault, parallel bars and horizontal bar. Rope climbing was also included in several early modern Olympics, and club swinging only at the 1904 Olympics. For women athletes there have always been four events: vault, uneven bars, balance beam and floor exercise.

In artistic gymnastics athletes perform short routines, ranging from approximately 30 to 90 seconds. Competitions use the New Life scoring rule, which was introduced in 1989. Under New Life, marks from one session do not carry over to the next. In other words, a gymnast's performance in team finals does not affect his or her scores in the all-around finals or event finals; he or she starts with a clean slate.

Only women athletes compete in rhythmic events which combine elements of ballet, gymnastics, dance, and working with pieces of kit. The sport involves the performance of five separate routines with the use of five different pieces of kit: ball, ribbon, hoop, clubs, and rope. There is a much greater emphasis on the grace and beauty of the routine rather than on the acrobatic aspects. There are events for individuals and also group routines performed by five women using five different pieces of kit.

Experiment 3 – Olympic Sports

The modern Olympic Games include 33 sports, 30 disciplines and nearly 400 events. For example, aquatics is a sport at the Summer Games which includes six disciplines. These are swimming, synchronised swimming, diving, water polo, open water swimming and high diving. Each discipline is then divided into a number of different events. For example, swimming currently involves 17 events each for men and women. The Summer Olympics include 26 sports, while the Winter Olympics feature15 sports.

The Summer Games always include athletics, swimming, fencing, and gymnastics. The Winter Games include cross-country and downhill skiing, figure skating, ice hockey, ski jumping, and speed skating. The Olympics Games usually also include demonstration sports which are typical of the host country. For example baseball was a demonstration sport at the 1984 Los Angeles Games. Some Olympic sports, such as basketball and volleyball, first appeared as demonstration sports and then became full Olympic sports.

On the other hand, previous Olympic Games included sports which are no longer part of the Games, for example polo and tug of war. In the early days of the modern Games, the local organisers were able to decide which sports to include, until the International Olympic Committee (IOC) took control of the programme in 1924.

These sports, known as discontinued sports, were usually dropped because of lack of interest. In addition, they may have been professional sports at a time when the Olympic Games were only for amateurs. Several such sports, such as archery and tennis, have since been readmitted to the Olympic programme. Curling was an official sport in the Winter Games until 1924, was then discontinued, but was reinstated as an Olympic sport in 1998.

Experimental 4 – Olympic Symbols

The Olympic Movement uses symbols to represent the ideals embodied in the Olympic Charter. The main symbol, the “Olympic Rings”, is composed of five interlocking rings coloured blue, yellow, black, green and red on a white field. The colours represent the flags of all the nations that were competing in the Games at the time of the introduction of the Olympic Rings.

The founder of the modern Olympic Games originally designed the rings in 1912. They were adopted in 1914 but the Games were suspended due to the hostilities of World War 1. The rings were officially adopted when the Games resumed after the war and used for the first time at the 1920 Summer Olympics. The popularity of the rings accelerated during the lead-up to the 1936 Summer Olympics in Berlin.

Months before each Games, another symbol, the “Olympic flame” is lit in Olympia, Greece. This ceremony reflects ancient Greek rituals. A female performer, acting as a priestess, ignites a torch by placing it inside a parabolic mirror that focuses the sun's rays. This then lights the torch of the first bearer of the flame in the Olympic torch relay, that will carry the flame to the Olympic stadium of the host city.

Another symbol is the “Olympic mascot”. This is an animal or human representation of the cultural heritage of the host country, and was introduced in 1968. It has played a particularly important part in the Games identity promotion since the 1980 Olympics, when the Russian bear cub, Misha, achieved international stardom. The mascot of the Summer Olympics in London in 2012 was christened Wenlock after the town of Much Wenlock which gave inspiration to the founder of the Olympic Games.

Experimental 5 – The Olympic Games Opening Ceremony

The opening ceremony of the Olympic Games includes a number of activities and takes place before the sporting events start. Although there was an Opening Ceremony at the first modern Games in 1896 in Athens, it was very simple compared to today's ceremonies. Some of the elements in the Modern Olympic Games traditions were gradually established over time through a series of adaptations from the early Games.

The opening ceremony starts with the raising of the host country's flag and a performance of its national anthem. Then the host nation presents displays of music, singing, dance, and theatre representing its culture. Usually, the content of these displays is kept secret until the last minute. The opening ceremony at the Beijing Games reportedly cost US$100 million, with much of the cost incurred in the artistic portion of the ceremony.

After the artistic portion of the opening ceremony, the athletes parade into the stadium grouped by nation. Greece is always the first nation to enter stadium in order to honour the origins of the Olympics Games in Greece. Nations then enter the stadium alphabetically according to the host country's chosen language, with the host country's athletes being the last to enter. Opening speeches are given to indicate the formal opening of the Games.

Finally, the Olympic torch is brought into the stadium and passed on until it reaches the final torch carrier. The final carrier is an Olympic athlete from the host nation, who lights the Olympic flame. The arrival of the flame starts another celebration of the Games. As doves are the symbol of peace, a release of doves follows the lighting of the Olympic flame.

Experimental 6 – The Olympic Games Closing Ceremony

The closing ceremony of the Olympic Games is organised on the last day of the Games after all sporting events have concluded. It is held in the Olympic stadium, normally after the last events. The first element is the entrance of the head of state of the host country, accompanied by the President of the International Olympic Committee (IOC) and the President of the Organising Committee for the Olympic Games (OCOG).

The national anthem of the host country is then played and its flag raised in the Olympic stadium. After that, the flag of Greece and the flag of the country hosting the next Olympic Games are also raised. An athlete chosen by their National Olympic Committee generally carries each flag. This is immediately followed by the athletes' parade. All the athletes walk together in no particular order during this parade.

After the athletes enter the stadium, the final medals ceremony of the Games is held. The organising committee of the respective host city, after consulting with the IOC, determines which event will have its medals presented. During the Summer Olympics, this is usually the men's marathon. Traditionally, the men's marathon is held in the last hours of competition on the last day of the Olympics, and the race finishes just before the closing ceremony.

The presidents of the OCOG and the IOC make closing speeches and the Olympic flame is extinguished. The mayor of the organising city transfers a special Olympic flag to the president of the IOC, who then passes it on to the mayor of the next hosting city. The next host nation then also briefly introduces the next Games with artistic displays of dance and theatre representative of its culture.

Experimental 7 – The Winter Olympic Games

The Winter Olympic Games feature sports conducted on snow and ice. An earlier competition, the Nordic Games, were organised in 1901 and were held again in 1903 and 1905 and then every fourth year until 1926. The organisers of the Nordic Games attempted to have winter sports added to the main Olympic programme but were unsuccessful until the 1908 Summer Olympics.

Three years later, the Italian Count Usseaux proposed that the International Olympic Committee (IOC) stage some winter sports as part of the 1912 Summer Olympics in Stockholm, Sweden. The organisers opposed this idea because they wished to protect the integrity of the Nordic Games and were concerned about a lack of facilities. The idea was resurrected for the 1916 Games, which were to be held in Berlin.

A winter sports week with speed skating, figure skating, ice hockey and Nordic skiing was planned, but the 1916 Olympics Games were cancelled after the outbreak of war. The first Winter Olympics were held in Chamonix, France in 1924. They were then held every four years until 1936 when they were again interrupted by the outbreak of war. St. Moritz, Switzerland was selected to host the first post-war Games in 1948 because of its neutrality during the war. Twenty-eight countries competed but athletes from Germany and Japan were not invited.

The Winter Games have evolved since their inception. Sports have been added and some of them, such as Alpine skiing, luge, and snowboarding, have earned a permanent spot in the Olympic programme. Others, such as curling and bobsleigh, have been discontinued and later reintroduced, or have been permanently discontinued. This last category includes military patrol although the modern Winter Olympic sport of biathlon is descended from it.

Experimental 8 – The Summer Olympic Games

The Summer Olympic Games were first held in1896 and are usually held every four years. Due to the success of the Summer Games, the International Olympic Committee (IOC) also created the Winter Olympic Games. Since 1904, Olympic gold, silver and bronze medals are awarded to the three athletes who finish first, second and third. The only country which has won with at least one gold medal at every Summer Games is the United Kingdom.

The United States has hosted four Summer Olympics, more than any other nation. The United Kingdom has hosted three Summer Olympics, all held in London. The only Olympics held in the southern hemisphere so far have both been in Australia, Melbourne in 1956 and Sydney in 2000. This year, Rio de Janeiro will host the first Summer Games to be held in South America.

The 1896 Summer Olympics, officially known as the Games of the First Olympiad, were held in Athens, Greece. It was chosen as the host city during a congress organised by Pierre de Coubertin, the French historian and founder of the modern Olympics. The IOC was also established during this congress. Despite many obstacles and setbacks, the 1896 Olympics were regarded as a great success.

Entry rules for each of the sports are set by the International Sports Federations (ISFs) that govern that sport's international competition. For sports for individual competitors, competitors typically qualify through attaining a certain place in a major international event or on the ISF's ranking list. The general rule is that a maximum of three individual athletes may represent each nation per competition. For team events, countries most often qualify through qualifying tournaments on each continent.

Experimental 9 – The Youth Olympic Games

The Youth Olympic Games are organised by the International Olympic Committee (IOC) for 14 to 18 year old athletes. The Youth Games are held every four years consistent with the current Olympic Games format. This youth version of the Olympic Games was approved in 2007 with the intention of sharing the hosting costs between the IOC and the host city.

The idea for the Youth Games came from Johann Rosenzopf in 1998. The introduction of these Games came from global concerns about youth obesity in developed nations and the decline of youth participation in sports activities worldwide. In addition, they aim to foster participation in the main Games. The IOC was initially negative about the idea of a purely sporting event for young people and requested that the event be as much about cultural education and exchange as it was about sports.

Over 200 different countries and 3,600 athletes participated in the inaugural 2010 Youth Summer Olympics. Participants are placed into three categories, 14–15 year olds, 16–17 year olds, and 17–18 year olds. Qualification to participate is determined by the IOC and International Sport Federations (ISFs) for the various sports in the programme.

The sports contested at the Youth Games are the same as those scheduled for the main Games, but with some adaptations, and a more limited number of disciplines and events. Education and culture are also key components for the Youth Games. Not only does the cultural aspect apply to athletes and participants, but also to youth around the world and inhabitants of the host city and surrounding regions. To this end a Culture and Education Programme is featured at each Games.

Experimental 10 – The Paralympic Games

The Paralympic Games are the Olympic Games for athletes who have a range of physical, visual and mental disabilities. There are both Summer and Winter Paralympic Games. Since the 1988 Summer Games in Seoul, South Korea, the Paralympic Games have been held immediately after the main Games.

The Paralympics have expanded from a small gathering of war veterans in 1948 to become one of the world's largest sporting events. Disabled athletes strive for comparable treatment with non-disabled athletes, but there is still a large funding gap between the two groups. In 2003, three American disabled athletes sued the United States Olympic Committee for not providing the same funding for disabled athletes as it did for able bodied athletes. The case was lost, but in the five years it took to be decided, funding for disabled athletes nearly tripled.

Given the wide range of disabled athletes, there are numerous competition categories in which they participate. Their disabilities are broken down into ten types. These types are further broken down into special classifications, which vary from sport to sport. The classification system has led to claims of cheating by athletes who may attempt to overstate their disability.

Some athletes with disabilities did compete in the main Olympic Games before the Paralympics started. The first disabled competitor was the American gymnast George Eyser in 1904, who had an artificial leg. Disabled athletes have also competed at the main Olympic Games since the Paralympics started. The most famous case has been Oscar Pistorius, the South African sprinter. Pistorius had both his legs amputated below the knee and raced with two carbon fibre blades. He attempted to qualify for the 2008 Summer Olympics and did qualify for the 2012 Games.

Experimental 11 – Boycotts of the Olympic Games

Australia, France, Great Britain, and Switzerland are the only countries to have competed at every Olympic Games since their inception in 1896. While countries sometimes miss an Olympics due to a lack of qualified athletes, some choose to boycott a celebration of the Games for various reasons. The first boycott was of the 1936 Berlin Games by Ireland, because of Germany’s treatment of Jewish people.

There were three boycotts of the 1956 Olympics in Melbourne, Australia. The Netherlands, Spain, and Switzerland all refused to attend because of the repression of the Hungarian uprising by the Soviet Union. Cambodia, Egypt, Iraq, and Lebanon boycotted the Games because of the Suez Crisis. China boycotted the Games because Taiwan was allowed to compete in the Games as the "Republic of China".

In 1972 and 1976 a large number of African countries threatened the International Olympic Committee (IOC) with a boycott. They forced the IOC to ban South Africa and Rhodesia because of their segregationist regimes. New Zealand was also one of the African boycott targets, because its national rugby union team had toured apartheid-ruled South Africa. The IOC conceded in the first two cases, but refused to ban New Zealand.

In 1980 and 1984, the main opponents in the Cold War, the United States and the Soviet Union, boycotted each other's Games. The United States and 64 other countries boycotted the 1980 Olympics because of the Soviet invasion of Afghanistan. This boycott reduced the number of nations participating to 81, the lowest number since 1956. The Soviet Union and 15 other nations countered by boycotting the 1984 Olympics, contending that they could not guarantee the safety of their athletes.

Experimental 12 - Performance Enhancing Drugs at the Olympic Games

In the early 20th century, many Olympic athletes began using drugs to improve and increase their athletic abilities. In 1904, strychnine was given to gold medallist Thomas Hicks by his own coach. The only Olympic death linked to performance enhancing drugs occurred at the 1960 Rome games. A Danish cyclist, Knud Jensen, fell from his bicycle and later died. A coroner's inquiry found that he was under the influence of amphetamines.

The first Olympic athlete to test positive for the use of performance enhancing drugs was Hans-Gunnar Liljenwall, a Swedish pentathlete at the 1968 Summer Olympics, who lost his bronze medal for alcohol use. One of the most publicised drugs related disqualifications occurred after the 1988 Summer Olympics when Canadian sprinter, Ben Johnson tested positive for stanozolol. His gold medal was later stripped and awarded to the American runner-up Carl Lewis.

In the late 1990s, the International Olympic Committee (IOC) took the initiative in a more organised battle against drug use, by forming the World Anti-Doping Agency (WADA) in 1999. There was a sharp increase in positive drug tests at the 2000 Summer Olympics and 2002 Winter Olympics. Several medallists in weightlifting and cross-country skiing were disqualified because of drugs offences. During the 2006 Winter Olympics, only one athlete failed a drugs test.

The IOC established a drug-testing regimen (now known as the Olympic Standard) which has set the worldwide benchmark that other sporting organisations around the world attempt to emulate. During the Beijing games, 3,667 athletes were tested by WADA. Both urine and blood tests were used to detect banned substances. National Olympic Committees barred several athletes because of their failures in both tests.

APPENDIX C – ADULT READING QUESTIONNAIRE (ARQ)

Answer ALL questions. Please circle your response for each question.

Question	Responses				
Do you think you are a good reader?	Yes *0*	No *1*	Don't Know *0.5*		
Can you read quickly and easily?	Yes *0*	No *1*	Don't Know *0.5*		
How good is your spelling?	Good *0*	Average *1*	Poor *2*	Very Poor *3*	
In your job, how often do you read?	Never *4*	Rarely *3*	Sometimes *2*	Frequently *1*	Always *0*
Do you find it difficult to read words you haven't seen before (e.g. place names?)	Never *0*	Rarely *1*	Sometimes *2*	Frequently *3*	Always *4*
Do you find it difficult to read aloud?	Never *0*	Rarely *1*	Sometimes *2*	Frequently *3*	Always *4*
Do you find it difficult to find the right word to say?	Never *0*	Rarely *1*	Sometimes *2*	Frequently *3*	Always *4*
Do you ever confuse the names of things?	Never *0*	Rarely *1*	Sometimes *2*	Frequently *3*	Always *4*
Do you confuse left and right?	Never *0*	Rarely *1*	Sometimes *2*	Frequently *3*	Always *4*
Do you have problems with organisation or time management?	Never *0*	Rarely *1*	Sometimes *2*	Frequently *3*	Always *4*
How often do you write in everyday life?	Never *4*	Rarely *3*	Sometimes *2*	Frequently *1*	Always *0*

Dyslexia is difficulty with reading and writing in people who: • **Do OK in other aspects of life (their difficulty is mostly with reading and writing)** • **Have had the chance to learn to read, but have not been able to learn like others**				
Based on this, do you think you are dyslexic?	Yes *1*	No *0*	Maybe *0.5*	
How would you rate your difficulties?	No Difficulties *0*	Mild *1*	Moderate *2*	Severe *3*
Has anyone ever raised any concerns about your reading?	Yes *1*	No *0*		
Have you ever had a diagnosis of dyslexia?	Yes *1*	No *0*		
If Yes, by whom?				

Scale:

0 to 10	Non-Dyslexic
11 to 21	Mild Dyslexic
22 to 32	Moderate Dyslexic
33 to 43	Severe Dyslexic

APPENDIX D – INFORMED CONSENT FORM

Thank you for offering to participate in my PhD research. This study is to investigate the experience that people have when reading the Web with types of presentations.

At the beginning of the session you will be asked to complete a short questionnaire with demographic information. You will then be asked to undertake a series of short tasks on a website and your eye movements will be recorded with an eye-tracking device. Finally, when all of the tasks are accomplished, you will be asked to complete another short questionnaire and will be given an opportunity to ask any questions you would like about this research study. At the end of this session, you will receive an example of your eye gaze pattern.

All information received during this study will be treated confidentially, and any results will be published in way that protects the anonymity of our participants. If you have any questions during the session, please feel free to ask. Further, you may withdraw from the study at any time.

You will receive a £10 gift voucher (Amazon or Marks and Spencer) for your participation in this study.

Section A

I, ___________________________, voluntarily consent to participate in this study on user experience of websites. I have been briefed about the basic nature and purpose of the project and feel that I understand it.

I understand that all data gathered will be treated confidentially. I understand that my data will only be available in its original form to Ili Farhana Md Mahtar and Prof. Helen Petrie. I understand that I will not be identified when the data is shared, described or interpreted.

I also understand that I may withdraw at any point during the study.

___________________________ ____________

Signature of research participant Date

___________________________ ____________

Signature of researcher Date

Section B

I have been adequately debriefed. I was not forced to complete the study. All my questions have been answered. I have been compensated for my participation as agreed.

Your signature: ___________________________ Date:_________________________

APPENDIX E – PRE-STUDY QUESTIONNAIRE

Part 1 Use of the Web

The following questions are about your use of the Web.

1. For approximately how long have you been using the Web?
   ☐ Less than 6 months ☐ 6-12 months ☐ 1-3 years
   ☐ 4-6 years ☐ More than 6 years
2. How often do you use the Web per week?
   ☐ Never ☐ 1-5 hours ☐ 6-10 hours
   ☐ 11-20 hours ☐ More than 20 hours
3. How would you rate your level of computer expertise (circle one)?
   Not Expert *Expert*
4. How would you rate your level of Web expertise (tick one)?
   Not Expert *Expert*

Part 2 Personal Data

Please answer the following general questions about yourself (this information is confidential and do not write your name).

1. Age: ……… years old
2. Gender: ☐ Male ☐ Female
3. What is your current employment status:
   ☐ Working full time or part time
   ☐ Studying
   ☐ Not working or retired
4. If working, what is your job title: ………………………………………….
5. If studying, what is your current study level:
   ☐ Undergraduate student
   ☐ Master student
   ☐ PhD student
   ☐ Other (please specify) ………………….
6. If studying, what is your area of study: …………………………………..

APPENDIX F – POST-STUDY QUESTIONNAIRE (NON-DYSLEXIC) FOR STUDY 1

In the website, you experienced 12 combinations of font types and font sizes. Use the web page provided to view the combinations again.

1. For each combination, please circle your rating of how **easy or difficult to read** that combination of text was.

Example: *Very easy to read* *Very difficult to read*

Z

Combination *Very easy to read* *Very difficult to read*

A

B

C

D

E

F

G

H

I

J

K

L

2. Out of the 12 combinations of font types and font sizes, please circle the combination which you **prefer** most.

A B C D E F G H I J K L

Why did you choose that combination?

……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………

3. Combination I and J are in Lexie Readable font, while combination K and L are in OpenDyslexic font, both developed to help readers with dyslexia.

   a. Have you encountered these fonts before?

   OpenDyslexic: ☐ Yes ☐ No
   Lexie Readable: ☐ Yes ☐ No

   b. Would you be comfortable in receiving work from team mates/students/ colleagues using these fonts?

   OpenDyslexic: *Not Comfortable* |—|—|—|—|—|—| *Very Comfortable*

   Why have you given this rating?

   ……………………………………………………………………………………
   ……………………………………………………………………………………
   ……………………………………………………………………………………

   Lexie Readable: *Not Comfortable* |—|—|—|—|—|—| *Very Comfortable*

   Why have you given this rating?

   ……………………………………………………………………………………
   ……………………………………………………………………………………
   ……………………………………………………………………………………

4. Any comments you have about the font types or font sizes or the reading task.

……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………

Thank you for spending your time to complete these questions

APPENDIX G – POST-STUDY QUESTIONNAIRE (DYSLEXIC) FOR STUDY 1

In the website, you experienced 12 combinations of font types and font sizes. Use the web page provided to view the combinations again.

1. For each combination, please circle your rating of how **easy or difficult to read** that combination of text was.

Example: *Very easy to read* … *Very difficult to read*

Z

Combination	*Very easy to read* … *Very difficult to read*
A	
B	
C	
D	
E	
F	
G	
H	
I	
J	
K	
L	

2. Out of the 12 combinations of font types and font sizes, please circle the combination which you **prefer** most.

A B C D E F G H I J K L

Why did you choose that combination?

……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………

3. Combination I and J are in Lexie Readable font, while combination K and L are in OpenDyslexic font, both developed to help readers with dyslexia.

   a. Have you encountered these fonts before?

OpenDyslexic: ☐ Yes ☐ No
Lexie Readable: ☐ Yes ☐ No

   b. Do you think it is easier to read than standard fonts?

No Easier … *A lot Easier*

OpenDyslexic: |——|——|——|——|——|——|

Lexie Readable: |——|——|——|——|——|——|

   c. Would you be interested in using these fonts on your computer?

Not at all Interested … *Very Interested*

OpenDyslexic: |——|——|——|——|——|——|

Why have you given this rating?
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………

Not at all Interested … *Very Interested*

Lexie Readable: |——|——|——|——|——|——|

Why have you given this rating?
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………

d. Would you be comfortable in submitting work to teachers / colleagues using these fonts?

	Not Comfortable					*Very Comfortable*
OpenDyslexic:						

Why have you given this rating?
…………………………………………………………………………………
…………………………………………………………………………………
…………………………………………………………………………………

	Not Comfortable					*Very Comfortable*
Lexie Readable:						

Why have you given this rating?
…………………………………………………………………………………
…………………………………………………………………………………
…………………………………………………………………………………

4. Any comments you have about the font types or font sizes or the reading task.

…………………………………………………………………………………
…………………………………………………………………………………
…………………………………………………………………………………
…………………………………………………………………………………
…………………………………………………………………………………
…………………………………………………………………………………
…………………………………………………………………………………

Thank you for spending your time to complete these questions

APPENDIX H – SEQUENCE OF TEXTS FOR PRACTICE AND EXPERIMENT TASKS IN STUDY 1

P	Sequence of Texts															
	Practice			Experimental												
1	1	2		1	2	3	4	5	6		7	8	9	10	11	12
2	1	2		2	3	4	5	6	7		8	9	10	11	12	1
3	1	2		3	4	5	6	7	8		9	10	11	12	1	2
4	1	2		4	5	6	7	8	9		10	11	12	1	2	3
5	1	2		5	6	7	8	9	10		11	12	1	2	3	4
6	1	2	BREAK	6	7	8	9	10	11	BREAK	12	1	2	3	4	5
7	1	2		7	8	9	10	11	12		1	2	3	4	5	6
8	1	2		8	9	10	11	12	1		2	3	4	5	6	7
9	1	2		9	10	11	12	1	2		3	4	5	6	7	8
10	1	2		10	11	12	1	2	3		4	5	6	7	8	9
11	1	2		11	12	1	2	3	4		5	6	7	8	9	10
12	1	2		12	1	2	3	4	5		6	7	8	9	10	11

Note. 'P' denotes Participants.

APPENDIX I – DESCRIPTIVE STATISTICS FOR RATINGS OF EASE OF READING IN STUDY 1

Category	Combination of Typeface and Font Size	Group	*M*	*SD*	*Mdn*	*IQR*
Sans Serif	Arial Small	1	5.33	1.23	5.50	2.00
		2	4.75	1.42	5.00	2.00
		3	4.33	1.50	4.00	3.00
		Total	**4.81**	**1.41**	**4.83**	
	Arial Medium	1	5.75	1.14	6.00	2.00
		2	5.42	1.24	6.00	2.00
		3	5.83	0.84	6.00	1.00
		Total	**5.67**	**1.07**	**6.00**	
	Verdana Small	1	5.42	1.44	6.00	3.00
		2	4.92	1.31	5.00	2.00
		3	4.17	1.19	4.50	2.00
		Total	**4.83**	**1.38**	**5.17**	
	Verdana Medium	1	5.92	1.00	6.00	2.00
		2	5.33	1.07	5.00	2.00
		3	5.08	1.51	5.00	2.00
		Total	**5.44**	**1.23**	**5.33**	
Serif	Times New Roman Small	1	5.33	1.16	6.00	2.00
		2	4.50	1.24	4.00	2.00
		3	3.92	1.62	3.50	4.00
		Total	**4.58**	**1.44**	**4.50**	
	Times New Roman Medium	1	5.67	1.23	6.00	2.00
		2	5.00	0.85	5.00	1.00
		3	4.92	1.44	4.50	2.00
		Total	**5.19**	**1.22**	**5.17**	
	Georgia Small	1	5.58	1.08	5.50	2.00
		2	5.17	1.03	5.00	2.00
		3	4.67	0.99	5.00	1.00
		Total	**5.14**	**1.07**	**5.17**	
	Georgia Medium	1	5.17	1.53	5.00	3.00
		2	5.08	0.90	5.00	1.00
		3	5.25	1.42	5.00	2.00
		Total	**5.17**	**1.28**	**5.00**	
Dyslexia-Optimised	Lexie Readable Small	1	4.33	2.02	4.00	5.00
		2	3.42	1.78	3.00	3.00
		3	3.25	1.55	3.00	2.00
		Total	**3.67**	**1.81**	**3.33**	

	Lexie Readable Medium	1	5.42	1.44	5.50	2.00
		2	4.92	1.62	5.50	2.00
		3	4.33	1.37	4.50	3.00
		Total	**4.89**	**1.51**	**5.17**	
	OpenDyslexic Small	1	3.25	1.77	3.00	4.00
		2	3.00	1.81	2.50	3.00
		3	3.67	1.97	3.00	3.00
		Total	**3.31**	**1.82**	**2.83**	
	OpenDyslexic Medium	1	3.50	1.51	3.50	2.00
		2	3.50	1.83	3.00	3.00
		3	3.67	1.88	3.50	3.00
		Total	**3.56**	**1.70**	**3.33**	

APPENDIX J – SEQUENCE OF TEXTS CREATED USING REPLICATED LATIN SQUARE DESIGN (RLSD) IN STUDY 2

In order to create a sequence of texts where each text was counterbalanced across all line spacing-line length condition, a RLSD was used. RLSD is an extension of a Latin square design. In Latin square, different treatments were placed in a balance fashion within a square block (The Pennsylvania State University, 2018). For replicated Latin square, two blocking factors were used.

This study used two repetitions of a 4 × 4 Latin square. In each repetition, two blocking factors were used which consists of a block of text and a block of version. The square block of order of texts had four treatments of experimental texts (1, 2, 3, 4), while the square block of versions had four treatments of line spacing-line length combinations (A, B, C, D) as follows:

VERSION A.	line spacing of 1.5 and line length of 60 – 70cpl
VERSION B.	line spacing of 2.0 and line length of 80 – 90cpl
VERSION C.	line spacing of 1.5 and line length of 80 – 90cpl
VERSION D.	line spacing of 2.0 and line length of 60 – 70cpl

The blocks were superimposed to produce a new block of combination of sequence. This process was repeated for another set of two blocking factors. In total, two blocks of combination of sequence were created and then merged into one block of final sequence for this study. Blocking factors used in creating the sequence are as follows:

	Order of texts					Order of versions					Combination of text and version			
	Block 1					Block 2					Superimposed Block			
Repetition 1	1	3	4	2	+	A	B	C	D	=	1A	3B	4C	2D
	4	2	1	3		B	A	D	C		4B	2A	1D	3C
	2	4	3	1		C	D	A	B		2C	4D	3A	1B
	3	1	2	4		D	C	B	A		3D	1C	2B	4A
	Block 1					Block 2					Superimposed Block			
Repetition 2	4	1	3	2	+	C	A	D	B	=	4C	1A	3D	2B
	2	3	1	4		A	C	B	D		2A	3C	1B	4D
	1	4	2	3		D	B	C	A		1D	4B	2C	3A
	3	2	4	1		B	D	A	C		3B	2D	4A	1C

Both blocks were merged to form one final block

The final block of sequence consists of eight different sequence of tasks. For each sequence, participant completed four tasks. In each task, participant read a text with one of the versions of combination of line spacing and line length. For instance, referring to the table below, Participant 1 was given the first task with Experimental Text 1 with version A (line spacing of 1.5 and line length of 60 – 70cpl), second task with Experimental Text 3 with version B (line spacing of 2.0 and line length of 80 – 90cpl), third task with Experimental Text 4 with version C (line spacing of 1.5 and line length of 80 – 90cpl), and fourth task with Experimental Text 2 with version D (line spacing of 2.0 and line length of 60 – 70cpl). Sequence of tasks for all participants in each group is presented in table below:

Participants	Task 1	Task 2	Task 3	Task 4
1	1A	3B	4C	2D
2	4B	2A	1D	3C
3	2C	4D	3A	1B
4	3D	1C	2B	4A
5	4C	1A	3D	2B
6	2A	3C	1B	4D
7	1D	4B	2C	3A
8	3B	2D	4A	1C

APPENDIX K – SAMPLE OF TEXT, QUIZ AND DIFFICULTY ASSESSMENT FOR STUDY 2

Instructions:
Please read the text then answer the 4 questions about it and rate how difficult you thought it was. You need to do this part without looking back at the text.

In this point, you can look back at the text. Now, turn over the page and rate again how difficult you think the question is, when you can see exactly what it says in the text. Also, please indicate whether you already know the fact about the question. There are 4 texts, please answer the questions about all of them.

Many thanks for your help!

Ili

1. Avocado

The avocado is a tree which came originally from Central Mexico. The word avocado also refers to the fruit from the tree. The fruit is a large berry containing a single large seed. The fruit have green-skinned, fleshy bodies that are usually pear or egg shaped. Avocados now grow in tropical and Mediterranean climates throughout the world. To propagate the trees farmers take grafts, this ensures good quality and quantity of fruit.

The Spanish described avocados in early documents about their conquest of Mexico. In Peru, archaeologists have found remains of an avocado which has been carbon-dated to over 10,000 years ago. Avocados are usually served raw, though some varieties, including the common 'Hass' variety, can be cooked for a short time. Avocado is often used in the dip known as guacamole. In some parts of the world avocado is frequently used for milkshakes and occasionally added to ice cream and other desserts.

The Spaniards brought avocados back to Europe in 1601. Hans Sloane made the first written mention of avocado in English in a 1696 index of Jamaican plants. In the United States, the avocado arrived in both Florida and Hawaii in 1833. Before 1915, the avocado was commonly called the alligator pear in the United States. In 1915, the California Avocado Association introduced the then innovative term avocado to refer to the plant. This name came from the Spanish name *ahuacate*.

Please answer these questions WITHOUT looking at the text (circle one of the crosses or put an X over it).

1. The avocado tree originated in …
    a. Central Guatemala
    b. Central India
    c. Central Mexico

How difficult was this question?

Very difficult *Not at all difficult*
+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

2. The modern name avocado evolved from …
    a. ahuacate
    b. alligator pear
    c. aardvark

How difficult was this question?

Very difficult *Not at all difficult*
+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

3. A predictable quality and quantity of avocados is achieved by …
    a. growing the trees in humid conditions
    b. propagating the trees by grafting
    c. controlling the amount of fertilisers for the trees

How difficult was this question?

Very difficult *Not at all difficult*
+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

4. A 10, 000 years old avocado has been discovered in ...
    a. California
    b. Florida
    c. Peru

How difficult was this question?

Very difficult *Not at all difficult*
+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

Now have a look at the text, and rate how hard you think the questions were when you can consult the text.

1. The avocado tree originated in …

How difficult was this question?

Very difficult *Not at all difficult*
\+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text ☐ YES ☐ NO

2. The modern name avocado evolved from …

How difficult was this question?

Very difficult *Not at all difficult*
\+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text ☐ YES ☐ NO

3. A predictable quality and quantity of avocados is achieved by …

How difficult was this question?

Very difficult *Not at all difficult*
\+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text ☐ YES ☐ NO

4. A 10, 000 years old avocado has been discovered in ...

How difficult was this question?

Very difficult *Not at all difficult*
\+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text ☐ YES ☐ NO

2. Chennai

Chennai, formerly Madras, is the capital of the Indian state of Tamil Nadu. It is located on the Coromandel Coast of the Bay of Bengal. It is one of the biggest cultural, economic and educational centres in South India. According to the 2011 Indian census, it is the sixth-largest city in India. The Quality of Living Survey rated Chennai as the safest city in India.

Chennai is among the most visited Indian cities for foreign tourists. It is home to many museums, galleries, and other institutions, many of which are major tourist attractions as well as playing a research role. The city has one of the oldest museums and art galleries in India, the Government Museum and the National Art Gallery. Both of these institutions were established in the early18th century. There is also the Fort Museum, which has an important collection of objects from the era of the British raj.

Chennai has the third-largest expatriate population in India, estimated at over 100,000 foreigners living in the city in 2016. In 2015 the BBC named Chennai the "hottest" city both to visit and to live. National Geographic magazine has ranked Chennai's food as second best in the world. It was the only Indian city to feature on the magazine’s list. In October 2017, UNESCO added Chennai to the Creative Cities Network (UCCN) list in recognition of its rich musical tradition.

3. Kyoto

Kyoto means "Capital City" in Japanese. Located in the central part of the island of Honshu in Japan, it was the Imperial capital for more than one thousand years. Archaeological evidence suggests there has been human settlement in Kyoto as far back as the Stone Age. However, we know relatively little about human activity in the area before the 6th century, around which time the Emperor Tenmu established the Shimogamo Shrine in the Sayko ward.

The UNESCO World Heritage Site of Ancient Kyoto includes seventeen locations within the city and its immediate vicinity. Of the monuments, 13 are Buddhist temples, three are Shinto shrines, and one is a castle. Kyoto has many historic buildings, unlike other Japanese cities that lost buildings to foreign invasions and war. Although wars, fires, and earthquakes ravaged Kyoto during its eleven centuries as the imperial capital, the destruction of World War II did not affect it.

Kyoto is located in a valley, part of the Yamashiro Basin, in the eastern part of the mountainous region known as the Tamba highlands. The Yamashiro Basin is surrounded on three sides by mountains. The original city was arranged in accordance with traditional Chinese *feng shui* principles. The Imperial Palace faced south, the right sector of the city being on the west while the left sector is on the east. The streets in the modern-day areas still follow a grid pattern.

4. Lettuce

Lettuce is an annual plant of the daisy family. It is usually grown as a leaf vegetable, but sometimes for its stem and seeds. Lettuce is often used for salads, although it is also used in other kinds of food, such as soups, sandwiches or grilled. One variety, the *woju*, *celtuce* or asparagus lettuce, is grown for its stems. These are eaten either raw or cooked in Chinese cuisine. Lettuce was eaten mainly in Europe and North America, but by the late 20th century the consumption of lettuce had spread throughout the world.

Lettuce was first cultivated by the ancient Egyptians who turned it from a weed whose seeds were used to produce oil, into a food plant grown for its succulent leaves. Lettuce spread to the Greeks and Romans. The Romans gave it the name *lactuca*, which led to the English word lettuce. Lettuce appeared often in medieval writings, including several herbals which described medical benefits of the plant.

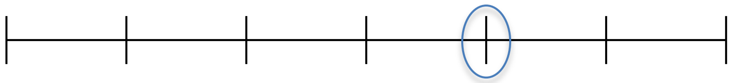
Generally grown as a hardy annual, lettuce is easily cultivated, although it requires relatively low temperatures to prevent it from flowering quickly. Lettuce is often plagued by numerous nutrient deficiencies, as well as insect and mammal pests, and fungal and bacterial diseases. Lettuce is a rich source of vitamin K and vitamin A, and a moderate source of folate and iron. Contaminated lettuce is often a source of bacterial, viral, and parasitic outbreaks in humans.

APPENDIX L – POST STUDY QUESTIONNAIRE FOR STUDY 2

In the study, you experienced 4 combinations of line length and line spacing. Use the web page provided to view the combinations again.

1. For each combination, please circle your rating of how **easy or difficult to read** that combination of text was. For example, if you thought a combination was moderately difficult to read, you would circle as below:

Example: *Very easy to read* … *Very difficult to read*

Z 

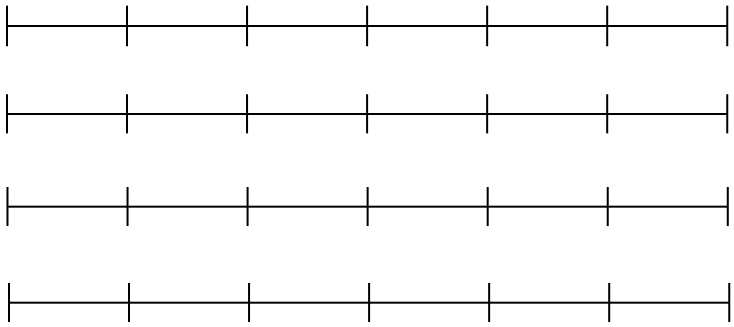
Combination *Very easy to read* … *Very difficult to read*

A

B

C

D



2. Out of the 4 combinations of line length and line spacing, please circle the combination which you prefer **most**.

A B C D

Why did you choose that combination?

………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………

3. Any comments you have about the line length or line spacing or the reading task.

………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………

Thank you for spending your time to complete these questions

APPENDIX M – QUESTIONS AND TASKS DIFFICULTY ASSESSMENT FOR STUDY 3

Please answer the following questions and rate how hard you think to find the answers.

1. How much does an Australian visitor need to pay for an electronic Travel Authorization (eTA)?
    a. $7
    b. $100
    c. US$6

    How difficult was it to find the answer to this question on the website?

    Very difficult *Not at all difficult*
    \+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

    I knew the answer before reading the text □ Yes □ No

2. When did the current 11-point maple leaf first appear on the Canada flag?
    a. 1836
    b. 1920
    c. 1965

    How difficult was it to find the answer to this question on the website?

    Very difficult *Not at all difficult*
    \+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

    I knew the answer before reading the text □ Yes □ No

3. When is the official season for camping?
    a. May-September
    b. July-August
    c. June-September

    How difficult was it to find the answer to this question on the website?

    Very difficult *Not at all difficult*
    \+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

    I knew the answer before reading the text □ Yes □ No

4. What kind of mountain biking will you experience in Vancouver's North Shore area?
    a. Gentle riverside trail
    b. Narrower and steeper trestles
    c. Converted railway barrels across wooden trestle bridges

How difficult was it to find the answer to this question on the website?

Very difficult *Not at all difficult*

+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text □ Yes □ No

5. What is the current capital city of the Yukon?
    a. Klondike
    b. Whitehorse
    c. Dawson City

How difficult was it to find the answer to this question on the website?

Very difficult *Not at all difficult*

+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text □ Yes □ No

6. What Canadian event celebrates outer space?
    a. Stratford Festival
    b. Festival Acadien
    c. Dark Sky Festival

How difficult was it to find the answer to this question on the website?

Very difficult *Not at all difficult*

+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text □ Yes □ No

7. Why do immigrants choose to stay in Ontario?
    a. It has well-established immigrant support services
    b. It has a vibrant multicultural neighbourhood
    c. Its spectacular scenery and mild climate

How difficult was it to find the answer to this question on the website?

Very difficult *Not at all difficult*

+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text □ Yes □ No

8. How many official border crossings are there along the US-Canadian border?
    a. Approximately 20
    b. Approximately 25
    c. Approximately 30

How difficult was it to find the answer to this question on the website?

Very difficult *Not at all difficult*

+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text □ Yes □ No

9. Confederation Bridge links Prince Edward Island with _________.
    a. New Brunswick
    b. Nova Scotia
    c. Yukon Territory

How difficult was it to find the answer to this question on the website?

Very difficult *Not at all difficult*

+ ------- + ------- + ------- + ------- + ------- + ------- + ------- + ------- +

I knew the answer before reading the text □ Yes □ No

APPENDIX N – FIRST POST-STUDY QUESTIONNAIRE FOR STUDY 3

In the study, you experienced a particular organisation of web navigation menus. The menu items were all at the top (and on the left, if you had any) of the page. Use the web site provided to view the web navigation menu organisation again and answer these questions.

1. Based on the web navigation menu organisation you had, please circle your rating for the following categories. For example, if you thought the menu organisation was moderately difficult to read, you would circle as below:

Example: *Very easy* *Very difficult*

Category

Very easy *Very difficult*

Ease of Use

Ease of Remembering

Ease of Learning

2. Please describe your strategy in general when navigating on websites, particularly how you use the menus.

…………………………………………………………………………………………
…………………………………………………………………………………………
…………………………………………………………………………………………
…………………………………………………………………………………………
…………………………………………………………………………………………
…………………………………………………………………………………………

3. Any comments you have about the organisation of web navigation menu.

………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………
………………………………………………………………………………………

Thank you for spending your time to complete these questions

APPENDIX O – SECOND POST-STUDY QUESTIONNAIRE FOR STUDY 3

With other organisation of web navigation menus, please rate for the following categories.

Ease of Use

Very easy … *Very difficult*

Website UV

Website UD

Website FV

Website FD

Ease of Remembering

Very easy … *Very difficult*

Website UV

Website UD

Website FV

Website FD

Ease of Learning

	Very easy						*Very difficult*
Website UV							
Website UD							
Website FV							
Website FD							

Any comments regarding the organisation of navigation menus.

……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………

Thank you for spending your time to complete these questions

APPENDIX P – SEQUENCE OF QUESTIONS FOR EXPERIMENT TASKS IN STUDY 3

Participants	Task 1	Task 2	Task 3	Task 4
1	Question 1	Question 2	Question 3	Question 4
2	Question 2	Question 3	Question 4	Question 1
3	Question 3	Question 4	Question 2	Question 1
4	Question 4	Question 1	Question 2	Question 3

In each group, two participants had the same sequence of questions in the experimental task.

REFERENCES

Aaltonen, A., Hyrskykari, A., & Räihä, K.-J. (1998). *101 spots, or how do users read menus?* Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.

Aderinokun, I. (2015). CSS Font Sizing. Retrieved 2 December 2016, from https://bitsofco.de/css-font-sizing/

Al-Wabil, A., Zaphiris, P., & Wilson, S. (2006). Web design for dyslexics: accessibility of arabic content *Computers Helping People with Special Needs* (pp. 817-822): Springer.

Al-Wabil, A., Zaphiris, P., & Wilson, S. (2007). Web navigation for individuals with dyslexia: an exploratory study *Universal Acess in Human Computer Interaction. Coping with Diversity* (pp. 593-602): Springer.

Al-Wabil, A., Zaphiris, P., & Wilson, S. (2008). *Examining visual attention of dyslexics on web navigation structures with eye tracking.* Paper presented at the Innovations in Information Technology, 2008. IIT 2008. International Conference on.

Andrews, E. (2018). Who invented the internet? - HISTORY. Retrieved 26th Dec 2018, from https://www.history.com/news/who-invented-the-internet

APEDA-France. (2010). Qu'est-ce que la dyslexie ? Retrieved 4th Nov 2018, from http://www.apeda-france.com/spip.php?article28

Armstrong, J. S. (1980). Unintelligible management research and academic prestige. *Interfaces, 10*(2), 80-86.

Bates, K. (2004). Lexia Readable: K-Type.

BDA Technology. (2015a). Spell Checkers. Retrieved 24 July 2015, from http://bdatech.org/what-technology/spell-checkers/

BDA Technology. (2015b). Text to Speech. Retrieved 24 July 2015, from http://bdatech.org/what-technology/text-to-speech/

BDA Technology. (2016). Typefaces for Dyslexia. Retrieved 2 June 2015, from http://bdatech.org/what-technology/typefaces-for-dyslexia/

Bernard, J.-B., Anne-Catherine, S., & Eric, C. (2007). Page mode reading with simulated scotomas: A modest effect of interline spacing on reading speed. *Vision research, 47*(28), 3447-3459.

Bernard, M. L., Fernandez, M., Hull, S., & Chaparro, B. S. (2003). *The Effects of Line Length on Children and Adults' Perceived and Actual Online Reading Performance.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Bernard, M. L., Hamblin, C. J., & Chaparro, B. S. (2003). *Comparing cascading and indexed menu designs for differences in performance and preference.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Blanca, M., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema, 29*(4), 552-557.

Boada, R., Willcutt, E. G., & Pennington, B. F. (2012). Understanding the comorbidity between dyslexia and attention-deficit/hyperactivity disorder. *Topics in Language Disorders, 32*(3), 264-284.

Boer, C. (Writer). (2011). How a typeface can help people [Video file]. In TEDxDubai (Producer).

British Dyslexia Association. (2012). Dyslexia style guide, January 2012.

British Dyslexia Association. (2018a). About the British Dyslexia Association. Retrieved 3rd November 2018, from https://www.bdadyslexia.org.uk/about

British Dyslexia Association. (2018b). Dyslexia Style Guide 2018: Creating Dyslexia Friendly Content.

Brunswick, N. (2010). Unimpaired reading development and dyslexia across different languages *Reading and dyslexia in different orthographies* (pp. 149-172): Psychology Press.

Burrell, A., & Sodan, A. C. (2006). *Web interface navigation design: which style of navigation-link menus do users prefer?* Paper presented at the 22nd International Conference on Data Engineering Workshops (ICDEW'06).

Bylinskii, Z., Borkin, M. A., Kim, N. W., Pfister, H., & Oliva, A. (2015). *Eye fixation metrics for large scale evaluation and comparison of information visualizations.* Paper presented at the Workshop on Eye Tracking and Visualization.

Chandler, S. B. (2001). *Comparing the legibility and comprehension of type size, font selection and rendering technology of onscreen type.* Virginia Tech.

Chatrangsan, M. (In Preparation). *Tablet computers for older people.* University of York.

Chen, M. C., Anderson, J. R., & Sohn, M. H. (2001). *What can a mouse cursor tell us more?: correlation of eye/mouse movements on web browsing.* Paper presented at the CHI'01 extended abstracts on Human factors in computing systems.

Chung, S. T. (2004). Reading speed benefits from increased vertical word spacing in normal peripheral vision. *Optometry and vision science: official publication of the American Academy of Optometry, 81*(7), 525.

Coe, R. (2002). It's the effect size, stupid: What effect size is and why it is important.

Colmer, R. (2019a). Readability and the Gunning Fog Index. Retrieved 20th Nov 2018, from https://readable.io/blog/the-gunning-fog-index/

Colmer, R. (2019b). The Flesch Reading Ease and Flesch-Kincaid Grade Level. Retrieved 20th Nov 2018, from https://readable.io/blog/the-flesch-reading-ease-and-flesch-kincaid-grade-level/

Coltheart, M. (2005). Modeling reading: The dual-route approach. *The science of reading: A handbook*, 6-23.

Cortiella, C., & Horowitz, S. H. (2014). The state of learning disabilities: Facts, trends and emerging issues. *New York: National center for learning disabilities*, 2-45.

Cotugna, N., Vickery, C. E., & Carpenter-Haefele, K. M. (2005). Evaluation of literacy level of patient education pages in health-related journals. *Journal of community health, 30*(3), 213-219.

Davey, B. (1988). Factors affecting the difficulty of reading comprehension items for successful and unsuccessful readers. *The Journal of Experimental Education, 56*(2), 67-76.

Davis, R. D. (1992). 37 Common Characteristics of Dyslexia. Retrieved 4th Nov 2018, from http://www.dyslexia.com/library/symptoms.htm de Leeuw, R. (2010). *Special font for dyslexia?* , University of Twente.

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*. Paper presented at the Proceedings of the International Cross-Disciplinary Conference on Web Accessibility.

DeCator, D. D. (2015). Winsorizing and Transforming Data. Retrieved 6th Dec 2018, from https://www.ddecator.com/blog/2015/6/winsorizing-and-transforming-data

Department for Education and Skills. (2004). A Framework for Understanding Dyslexia Available from https://www.texthelp.com/media/39354/USAdultLiteracy.pdf

Dickinson, A., Gregor, P., & Dickinson, L. (2003). *SeeWord: Rethinking Interfaces, Insights from Word Processing Software for Dyslexic Readers*. Paper presented at the INTERACT.

Dickinson, A., Gregor, P., & Newell, A. F. (2002). *Ongoing investigation of the ways in which some of the problems encountered by some dyslexics can be alleviated using computer techniques*. Paper presented at the Proceedings of the fifth international ACM conference on Assistive technologies.

Dickman, E. (2017). Do We Need a New Definition of Dyslexia. Retrieved 4th Nov 2018, from https://dyslexiaida.org/do-we-need-a-new-definition-of-dyslexia/

Disability Rights Commission. (2004). The web: Access and inclusion for disabled people. *The Stationary Office*.

Draper, S. (2018). Effect size. Retrieved 21st Dec 2018, from http://www.psy.gla.ac.uk/~steve/best/effect.html

Duchnicky, R. L., & Kolers, P. A. (1983). Readability of text scrolled on visual display terminals as a function of window size. *Human Factors, 25*(6), 683-692.

Dyslexia ScotWest. (n.d.). Meares-Irlen Syndrome. Retrieved 6 July 2015, from http://www.dyslexiasw.com/advice/all-about-dyslexia/meares-irlen-syndrome

DyslexieFont.com. (2012). Research. Retrieved 20 June 2015, from http://www.dyslexiefont.com/en/dyslexia-font/research/

DyslexieFont.com. (n.d.). How the font was conceived. Retrieved 16 June 2015, from http://www.dyslexiefont.com/en/dyslexia-font/

Dyson, M. C. (2004). How physical text layout affects reading from screen. *Behaviour & Information Technology, 23*(6), 377-393.

Dyson, M. C., & Haselgrove, M. (2000). The effects of reading speed and reading patterns on the understanding of text read from screen. *Journal of Research in Reading, 23*(2), 210-223.

Dyson, M. C., & Haselgrove, M. (2001). The influence of reading speed and line length on the effectiveness of reading from screen. *International Journal of Human-Computer Studies, 54*(4), 585-612.

Dyson, M. C., & Kipping, G. J. (1998). The effects of line length and method of movement on patterns of reading from screen. *Visible Language, 32*(2), 150.

Elliott, J. G., & Gibbs, S. (2008). Does dyslexia exist? *Journal of Philosophy of Education, 42*(3‐4), 475-491.

Elliott, J. G., & Grigorenko, E. L. (2014). *The dyslexia debate*: Cambridge University Press.

Ellis, A. W. (2016). *Reading, writing and dyslexia (classic edition): a cognitive analysis*: Psychology Press.

Eraslan, S., Yesilada, Y., & Harper, S. (2015). Eye tracking scanpath analysis techniques on web pages: A survey, evaluation and comparison. *Journal of Eye Movement Research, 9*(1).

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.

Fang, X., & Holsapple, C. W. (2007). An empirical study of web site navigation structures' impacts on web site usability. *Decision Support Systems, 43*(2), 476-491.

Farnsworth, B. (2018). 10 Most Used Eye Tracking Metrics and Terms. Retrieved 23rd Dec 2018, from https://imotions.com/blog/7-terms-metrics-eye-tracking/

Faulkner, X., & Hayton, C. (2011). When left might not be right. *Journal of Usability Studies, 6*(4), 245-256.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*: sage.

Fitts, P. M., Jones, R. E., & Milton, J. L. (2005). Eye movements of aircraft pilots during instrument-landing approaches. *Ergonomics: Psychological mechanisms and models in ergonomics, 3*, 56.

Flesch, R. (1948). A new readability yardstick. *Journal of applied psychology, 32*(3), 221.

Fowler, S., & Stanwick, V. (2004). *Web application design handbook: Best practices for web-based software*: Morgan Kaufmann.

Freire, A. P., Petrie, H., & Power, C. (2011). *Empirical results from an evaluation of the accessibility of websites by dyslexic users.* Paper presented at the Proceedings of the Workshop on Accessible Design in the Digital World. Lisboa, Portugal: CEUR-WS. org.

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

Frith, U. (1999). Paradoxes in the definition of dyslexia. *Dyslexia, 5*(4), 192-214.

Gelderman, M. (1999). A short introduction to font characteristics. *TUGboat, 20*(2), 81-88.

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

Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of educational research, 42*(3), 237-288.

Gonzalez, A. (2015). OpenDyslexic. *Retrieved on December, 13*, 2015.

Goulandris, N. E. (2003). *Dyslexia in different languages: Cross-linguistic comparisons*: Whurr Publishers.

Gregor, P., & Dickinson, A. (2007). Cognitive difficulties and access to information systems: an interaction design perspective. *Universal Access in the Information Society, 5*(4), 393-400.

Gregor, P., Dickinson, A., Macaffer, A., & Andreasen, P. (2003). SeeWord - A personal word processing environment for dyslexic computer users. *British Journal of Educational Technology, 34*(3), 341-355.

Gregor, P., & Newell, A. F. (2000). *An empirical investigation of ways in which some of the problems encountered by some dyslexics may be alleviated using computer techniques*. Paper presented at the Proceedings of the fourth international ACM conference on Assistive technologies.

Grigorovich-Barsky, M., & Belson, D. S. (2013). The Effects of Fonts on Reading Performance for Those with Dyslexia: A Quasi-Experimental Study.

Hanley, J. R. (2017). Is There Just One Dyslexic Reader? Evidence for the Existence of Distinct Dyslexic Sub-Groups. *Current Developmental Disorders Reports, 4*(4), 101-107.

Harley, L., Kline, K., Price, C., Jones, A., Mosley, S., Farmer, S., et al. (2013). A Web Based Voting Application Study of Fonts for Voters with Dyslexia. *The Information Technology and Innovation Foundation Accessible Voting Technology Initiative Working Paper Series*, 1-30.

Harrysson, B., Svensk, A., & Johansson, G. (2004). How people with developmental disabilities navigate the Internet. *British Journal of Special Education, 31*(3), 138-142.

Henry, M. L., Beeson, P. M., Stark, A. J., & Rapcsak, S. Z. (2007). The role of left perisylvian cortical regions in spelling. *Brain and language, 100*(1), 44-52.

Hillier, R. A. (2006). *A typeface for the adult dyslexic reader.* Anglia Ruskin University Cambridge.

Huang, J., White, R., & Buscher, G. (2012). *User see, user point: gaze and cursor alignment in web search.* Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(6), 1430.

Interagency Committee on Learning Disabilities. (1987). *Learning Disabilities: A Report to U.S. Congress*. Washington, DC.

International Dyslexia Association. (2002). Definition of Dyslexia. Retrieved 4th Nov 2018, from http://eida.org/definition-of-dyslexia/

International Dyslexia Association. (n.d.). Dyslexia Basics. Retrieved 19th April 2019, from https://dyslexiaida.org/dyslexia-basics/

Internet Live Statistics. (2018). Internet Users. Retrieved 3rd Nov 2018, from http://www.internetlivestats.com/internet-users/

Irlen UK. (n.d.). Understanding Irlen Syndrome. Retrieved 6 July 2015, from http://www.irlen.org.uk

Ismail, R., & Jaafar, A. (2011). *Interactive screen-based design for dyslexic children.* Paper presented at the User Science and Engineering (i-USEr), 2011 International Conference on.

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

K-Type. (2008). Lexia Readable. Retrieved 2 June 2015, from http://www.k-type.com/lexia-readable/

Kalbach, J. (2007). *Designing Web Navigation*: O'Reilly Media.

Kalbach, J., & Bosenick, T. (2003). Web page layout: A comparison between left-and right-justified site navigation menus. *Journal of Digital Information, 4*(1).

Kamollimsakul, S. (2014). *Web Design Guidelines for Text Presentation for Older People: Empirical Evidence from Thailand and the UK.* University of York.

Kanvinde, G., Rello, L., & Baeza-Yates, R. (2012). *IDEAL: a dyslexic-friendly ebook reader.* Paper presented at the Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility.

Kaspar, K., Wehlitz, T., von Knobelsdorff, S., Wulf, T., & von Saldern, M. A. O. (2015). A matter of font type: The effect of serifs on the evaluation of scientific abstracts. *International Journal of Psychology, 50*(5), 372-378.

Katusic, S. K., Colligan, R. C., Barbaresi, W. J., Schaid, D. J., & Jacobsen, S. J. (2001). *Incidence of reading disability in a population-based birth cohort, 1976–1982, Rochester, Minn.* Paper presented at the Mayo Clinic Proceedings.

Katzir, T., Hershko, S., & Halamish, V. (2013). The Effect of Font Size on Reading Comprehension on Second and Fifth Grade Children: Bigger Is Not Always Better. *PloS one, 8*(9), e74061.

Kessler, T. (2010). Understanding relative vs. absolute font sizes on Web pages. Retrieved 19th Nov 2018, from https://www.cnet.com/news/understanding-relative-vs-absolute-font-sizes-on-web-pages/

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel.

Kinsman, P. (1997). Teaching Phoneme-grapheme Correspondence Rules to Children with Learning Difficulties: An Implementation of the Fernald Method.

Klare, G. R. (1984). Readability. *Handbook of reading research, 1*, 681-744.

Krithikadatta, J. (2014). Normal distribution. *Journal of conservative dentistry: JCD, 17*(1), 96.

Kuster, S. M., van Weerdenburg, M., Gompel, M., & Bosman, A. M. (2017). Dyslexie font does not benefit reading in children with or without dyslexia. *Annals of dyslexia*, 1-18.

Laerd Statistics. (2015). Three-way mixed ANOVA (with one between-subjects and two within-subjects factors) using SPSS Statistics. Retrieved 15th Nov 2018, from https://statistics.laerd.com/

Lawler, B. P. (2006). *The Official Adobe Print Publishing Guide* (2 ed.): Adobe Press.

Leeuw, R. (2010). *Special font for dyslexia?* , University of Twente.

Legge, G. E., & Bigelow, C. A. (2011). Does print size matter for reading? A review of findings from vision science and typography. *Journal of Vision, 11*(5), 8.

Leuthold, S., Schmutz, P., Bargas-Avila, J. A., Tuch, A. N., & Opwis, K. (2011). Vertical versus dynamic menus on the world wide web: Eye tracking study measuring the influence of menu design and task complexity on user performance and subjective preference. *Computers in human behavior, 27*(1), 459-472.

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision research, 48*(5), 635-654.

Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research, 28*(4), 612-625.

Lonely Planet. (2017). *Canada* (13 ed.): Lonely Planet Global Limited.

Lukeš, D. (2015). *Dyslexia Friendly Reader.* Paper presented at the 6th International Conference on Information, Intelligence, Systems and Applications (IISA), Corfu, Greece.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of dyslexia, 53*(1), 1-14.

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

Malpas, M. (2012). *Adults and Dyslexia, 40 years on...* British Dyslexia Association.

Marinus, E., Mostard, M., Segers, E., Schubert, T. M., Madelaine, A., & Wheldall, K. (2016). A Special Font for People with Dyslexia: Does it Work and, if so, why? *Dyslexia, 22*(3), 233-244.

Martelli, M., Di Filippo, G., Spinelli, D., & Zoccolotti, P. (2009). Crowding, reading, and developmental dyslexia. *Journal of Vision, 9*(4), 14.

McCarthy, J. D., Sasse, M. A., & Riegelsberger, J. (2004). Could I have the menu please? An eye tracking study of design conventions *People and computers XVII—Designing for society* (pp. 401-414): Springer.

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.

McLoughlin, D., Fitzgibbon, G., & Young, V. (1994). *Adult dyslexia: assessment, counselling and training*. London: Whurr Publisher Ltd.

Meindertsma, J. (2016). The power of typefaces: Dyslexie. *TXT, 2016*, 4.

Michail, K. (2010). *Dyslexia: the experiences of university students with dyslexia.* University of Birmingham.

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.

Moll, K., & Jones, M. (2013). Naming fluency in dyslexic and nondyslexic readers: Differential effects of visual crowding in foveal, parafoveal, and peripheral vision. *The Quarterly Journal of Experimental Psychology, 66*(11), 2085-2091.

Montani, V., Facoetti, A., & Zorzi, M. (2014). The effect of decreased interletter spacing on orthographic processing. *Psychonomic bulletin & review, 22*(3), 824-832.

Morris, M. R., Fourney, A., Ali, A., & Vonessen, L. (2018). *Understanding the Needs of Searchers with Dyslexia.* Paper presented at the Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.

Murano, P., & Lomas, T. J. (2015). Menu Positioning on Web Pages. Does it Matter?

Murano, P., & Oenga, K. K. (2012). The Impact on Effectiveness and User Satisfaction of Menu Positioning on Web Pages. *positions, 3*(9).

Murano, P., & Sander, M. (2016). User interface menu design performance and user preferences: a review and ways forward. *Int. J. Adv. Comput. Sci. Appl, 7*(4), 355-361.

National Health Service. (2014). Dyslexia. Retrieved 4th Nov 2018, from http://www.nhs.uk/Conditions/Dyslexia/Pages/Introduction.aspx

National Research Council, Division of Behavioral and Social Sciences and Education, Board of Behavioral, C. a. S. S., & Impairments, C. o. D. D. f. I. w. V. (2002). *Visual Impairments: Determining Eligibility for Social Security Benefits*. Washington DC: National Academies Press.

Ndombo, D. M., Ojo, S., Osunmakinde, I. O., & Phasha, N. (2013). *Evidence of dyslexia in special schools in south africa and a proposal of assistive technology*. Paper presented at the IST-Africa Conference and Exhibition (IST-Africa), 2013.

Nielsen, J. (1999). *Designing web usability: The practice of simplicity*: New Riders Publishing.

O'Brien, B. A., Mansfield, J. S., & Legge, G. E. (2005). The effect of print size on reading speed in dyslexia. *Journal of Research in Reading, 28*(3), 332-349.

OpenDyslexic. (n.d.). About. Retrieved 12 June 2015, from http://opendyslexic.org/about/

Patsula, P. J., Detenber, B. H., & Theng, Y.-L. (2010). Structure processing of web-based menus. *Intl. Journal of Human–Computer Interaction, 26*(7), 675-702.

Pauc, R. (2005). Comorbidity of dyslexia, dyspraxia, attention deficit disorder (ADD), attention deficit hyperactive disorder (ADHD), obsessive compulsive disorder (OCD) and Tourette's syndrome in children: A prospective epidemiological study. *Clinical chiropractic, 8*(4), 189-198.

Perea, M., & Gomez, P. (2012). Subtle increases in interletter spacing facilitate the encoding of words during normal reading.

Perea, M., Panadero, V., Moret-Tatay, C., & Gómez, P. (2012). The effects of inter-letter spacing in visual-word recognition: Evidence with young normal readers and developmental dyslexics. *Learning and Instruction, 22*(6), 420-430.

Pernice, K., & Nielsen, J. (2009). How to conduct eyetracking studies. *Nielsen Norman Group*.

Pijpker, C. (2013). *Reading performance of dyslexics with a special font and a colored background.* University of Twente.

Pittsley, K. A., & Memmott, S. (2012). Improving independent student navigation of complex educational web sites: An analysis of two navigation design changes in LibGuides. *Information Technology and Libraries, 31*(3), 52-64.

Precision Vision. (n.d.). Original Series ETDRS Chart. Retrieved 21 June 2018, from https://www.precision-vision.com/product/4meteroriginalseriesetdrschart1/

Puerta Melguizo, M. C., Vidya, U., & van Oostendorp, H. (2012). Seeking information online: the influence of menu type, navigation path complexity and spatial ability on information gathering tasks. *Behaviour & Information Technology, 31*(1), 59-70.

Rabinowitz, T. (2015). *Exploring typography*: Cengage Learning.

Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly*, 29-53.

Ramus, F. (2001). Dyslexia: Talk of two theories. *Nature, 412*(6845), 393.

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., et al. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain, 126*(4), 841-865.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin, 124*(3), 372.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr, C. (2012). *Psychology of reading*: Psychology Press.

Reid, G. (2011). *Dyslexia*: A&C Black.

Reid, G., & Green, S. (1996). A definition of dyslexia. *Edinburgh, Moray House Centre for the*.

Rello, L. (2012). DysWebxia: a model to improve accessibility of the textual web for dyslexic users. *ACM SIGACCESS Accessibility and Computing*(102), 41-44.

Rello, L. (2014a). DysWebxia : A Text Accessibility Model for People with Dyslexia. In R. Baeza-Yates, H. Saggion & I. L. C. Universitat Pompeu Fabra. Departament De Tecnologies De La Informaci (Eds.).

Rello, L. (2014b). *DysWebxia: a text accessibility model for people with dyslexia.* Universitat Pompeu Fabra.

Rello, L., & Baeza-Yates, R. (2012). *Optimal Colors to Improve Readability for people with Dyslexia.* Paper presented at the Text Customization for Readability Symposium Online Symposium.

Rello, L., & Baeza-Yates, R. (2013). *Good fonts for dyslexia.* Paper presented at the Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility.

Rello, L., & Baeza-Yates, R. (2015). How to present more readable text for people with dyslexia. *Universal Access in the Information Society*, 1-21.

Rello, L., & Baeza-Yates, R. (2016). The Effect of Font Type on Screen Readability by People with Dyslexia. *ACM Transactions on Accessible Computing (TACCESS), 8*(4), 15.

Rello, L., Baeza-Yates, R., Saggion, H., Bayarri, C., & Barbosa, S. D. (2013). *An iOS reader for people with dyslexia.* Paper presented at the Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility.

Rello, L., Baeza-Yates, R. A., & Saggion, H. (2013). *The Impact of Lexical Simplification by Verbal Paraphrases for People with and without Dyslexia.* Paper presented at the CICLing (2).

Rello, L., & Barbosa, S. D. (2013). *Do People with Dyslexia Need Special Reading Software?* Paper presented at the INTERACT 2013 Cape Town, South Africa.

Rello, L., Bayarri, C., Gòrriz, A., Baeza-Yates, R., Gupta, S., Kanvinde, G., et al. (2013). *Dyswebxia 2.0!: more accessible text for people with dyslexia.* Paper presented at the Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility.

Rello, L., Bayarri, C., Otal, Y., & Pielot, M. (2014). *A Method to Improve the Spelling of Children with Dyslexia.* Paper presented at the ASSET'14, New York, USA.

Rello, L., & Bigham, J. P. (2017). *Good Background Colors for Readers: A Study of People with and without Dyslexia.* Paper presented at the Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility.

Rello, L., Kanvinde, G., & Baeza-Yates, R. (2012a). A Mobile Application for Displaying More Accessible eBooks for People with Dyslexia. *Procedia Computer Science, 14*, 226-233.

Rello, L., Kanvinde, G., & Baeza-Yates, R. (2012b). *Layout guidelines for web text and a web service to improve accessibility for dyslexics*. Paper presented at the Proceedings of the international cross-disciplinary conference on web accessibility.

Rello, L., Pielot, M., Marcos, M.-C., & Carlini, R. (2013). *Size matters (spacing not): 18 points for a dyslexic-friendly Wikipedia*. Paper presented at the Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility.

Riddick, B., Farmer, M., & Sterling, C. (1997). *Students and dyslexia: growing up with a specific learning difficulty*. London: Whurr.

Rose, J. (2009). Identifying and teaching children and young people with dyslexia and literacy difficulties: an independent report.

Schaeffer, K. (2008). CSS Font-Size: em vs. px vs. pt vs. percent. Retrieved 20th Nov 2018, from http://kyleschaeffer.com/development/css-font-size-em-vs-px-vs-pt-vs/

Schattka, K. I., Radach, R., & Huber, W. (2010). Eye movement correlates of acquired central dyslexia. *Neuropsychologia, 48*(10), 2959-2973.

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust? *Methodology*.

Schneps, M. H., Thomson, J. M., Chen, C., Sonnert, G., & Pomplun, M. (2013). E-readers are more effective than paper for some with dyslexia. *PloS one, 8*(9), e75634.

Schneps, M. H., Thomson, J. M., Sonnert, G., Pomplun, M., Chen, C., & Heffner-Wong, A. (2013). Shorter lines facilitate reading in those who struggle. *PloS one, 8*(8), e71161.

Schoonewelle, A. (2013). *Typographical support for dyslexics: The effect of textual alterations on dyslexic's test scores*. Vrije Universiteit Amsterdam.

Schulz, T. (2016). *Internal Validity in Experiments for Typefaces for People with Dyslexia*. Paper presented at the International Conference on Computers Helping People with Special Needs.

Selikowitz, M. (1998). *Dyslexia and other learning difficulties : the facts* (2nd ed.). Oxford ; New York: Oxford University Press.

Sevilla, J., Herrera, G., Martínez, B., & Alcantud, F. (2007). Web accessibility for individuals with cognitive deficits: A comparative study between an existing commercial web and its cognitively accessible equivalent. *ACM Transactions on Computer-Human Interaction (TOCHI), 14*(3), 12.

Shaikh, A. D. (2005). The effects of line length on reading online news. *Usability News, 7*(2), 1-4.

Shaywitz, S. E., & Shaywitz, B. A. (2005). Dyslexia (specific reading disability). *Biological psychiatry, 57*(11), 1301-1309.

Siegel, L. S. (2006). Perspectives on dyslexia. *Paediatrics & child health, 11*(9), 581-587.

Simmons, F., & Singleton, C. (2000). The reading comprehension abilities of dyslexic students in higher education. *Dyslexia, 6*(3), 178-192.

Sjoblom, A. M., Eaton, E., & Stagg, S. D. (2016). The effects of letter spacing and coloured overlays on reading speed and accuracy in adult dyslexia. *British Journal of Educational Psychology, 86*(4), 630-639.

Skottun, B. C., & Skoyles, J. R. (2012). Interletter spacing and dyslexia. *Proceedings of the National Academy of Sciences, 109*(44), E2958-E2958.

Snowling, M. J. (1998). Dyslexia as a phonological deficit: Evidence and implications. *Child Psychology and Psychiatry Review, 3*(1), 4-11.

Snowling, M. J. (2000). *Dyslexia* (2nd ed. ed.). Malden, MA: Blackwell Publishers.

Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia, 7*(1), 37-46.

Snowling, M. J. (2012). Changing concepts of dyslexia: nature, treatment and comorbidity. *Journal of Child Psychology and Psychiatry, 53*(9), e1-e3.

Snowling, M. J. (2013). Early identification and interventions for dyslexia: a contemporary view. *Journal of Research in Special Educational Needs, 13*(1), 7-14.

Snowling, M. J., Dawes, P., Nash, H., & Hulme, C. (2012). Validity of a Protocol for Adult Self‐Report of Dyslexia and Related Difficulties. *Dyslexia, 18*(1), 1-15.

Spencer, K. (2000). Is English a dyslexic language? *Dyslexia, 6*(2), 152-162.

Spinelli, D., De Luca, M., Judica, A., & Zoccolotti, P. (2002). Crowding effects on word identification in developmental dyslexia. *Cortex, 38*(2), 179-200.

Stanovich, K. E. (1996). Toward a more inclusive definition of dyslexia. *Dyslexia, 2*(3), 154-166.

Tan, L. H., Laird, A. R., Li, K., & Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta‐analysis. *Human brain mapping, 25*(1), 83-91.

Tavakoli, E., & Kheirzadeh, S. (2011). The Effect of Font Size on Reading Comprehension Skills: Scanning for Key Words and Reading for General Idea. *Theory and Practice in Language Studies*, 915.

The Pennsylvania State University. (2018). 4.4 - Replicated Latin Squares | STAT 503. Retrieved 10th Dec 2018, from https://onlinecourses.science.psu.edu/stat503/node/22/

Thomson, M. E., & Watkins, E. J. (1990). *Dyslexia : a teaching handbook*. London: Whurr.

Tinker, M. A., & Paterson, D. G. (1955). The effect of typographical variations upon eye movement in reading. *The Journal of Educational Research, 49*(3), 171-184.

Turner, M. (1997). *Psychological assessment of dyslexia*. London: Whurr.

van de Vrugt, J., & Ossen, A. (2012). *Dyslexie Regular Research 2012: Research on daily users, teachers and students*.

Vellutino, F., Fletcher, J., Snowling, M. J., & Scanlon, D. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of child psychology and psychiatry, 45*(1), 2-40.

Vidyasagar, T. R., & Pammer, K. (2010). Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. *Trends in cognitive sciences, 14*(2), 57-63.

W3C. (2000). Techniques for Accessibility Evaluation and Repair Tools. Retrieved 7 July 2015, from http://www.w3.org/TR/AERT

W3C. (2014). How to meet WCAG 2.0. Retrieved 28 July 2015, from http://www.w3.org/WAI/WCAG20/quickref/#qr-visual-audio-contrast-contrast

W3C. (2019). How to Meet WCAG 2.0 (Quick Reference). Retrieved 12th May 2019, from https://www.w3.org/WAI/WCAG21/quickref/

W3C WAI. (2018). Introduction to Web Accessibility. Retrieved 4th Nov 2018, from https://www.w3.org/WAI/fundamentals/accessibility-intro/

W3schools.com. (n.d.). HTML Color Names. July 2018, from https://www.w3schools.com/colors/colors_names.asp

W3Techs. (2015). Usage of content languages for websites. Retrieved 12 October 2015, from http://w3techs.com/technologies/overview/content_language/all

Washburn, E. K., Binks‐Cantrell, E. S., & Joshi, R. (2014). What do preservice teachers from the USA and the UK know about dyslexia? *Dyslexia, 20*(1), 1-18.

WebAIM. (2016). Introduction to Web Accessibility. Retrieved 4th Nov 2018, from http://webaim.org/intro/

Wery, J. J., & Diliberto, J. A. (2017). The effect of a specialized dyslexia font, OpenDyslexic, on reading rate and accuracy. *Annals of dyslexia, 67*(2), 114-127.

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in cognitive sciences, 15*(4), 160-168.

Williams, P., & Hennig, C. (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 Educational Needs, 15*(1), 25-36.

Willings, C. (2017). Environmental Print Comparisons. Retrieved 1 May 2018, from https://www.teachingvisuallyimpaired.com/print-comparisons.html

Wilson, J. R., & Corlett, N. (2005). *Evaluation of human work*: CRC press.

World Health Organization. (2015). Disability and Health. Retrieved 4th Nov 2018, from http://www.who.int/mediacentre/factsheets/fs352/en/

Wydell, T. N., & Butterworth, B. (1999). A case study of an English-Japanese bilingual with monolingual dyslexia. *Cognition, 70*(3), 273-305.

Yu, B. M., & Roh, S. Z. (2002). The effects of menu design on information‐seeking performance and user's attitude on the World Wide Web. *Journal of the American Society for Information Science and Technology, 53*(11), 923-933.

Zarach, V. (2002). Ten guidelines for improving accessibility for people with dyslexia. *CETIS University of Wales Bangor*.

Zeman, A., Milton, F., Smith, A., & Rylance, R. (2013). By heart an fMRI study of brain activation by poetry and prose. *Journal of Consciousness Studies, 20*(9-10), 132-158.

Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M., et al. (2012). Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences, 109*(28), 11455-11459.