Visual Constraints on Reading in Normal Vision, Low Vision and Dyslexia

A Dissertation
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Nilsu Atilgan

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisors: Sheng He and Gordon E Legge

December 2020

Acknowledgments

I feel incredibly fortunate to have had the opportunity to know and work with my wonderful advisors, Drs. Gordon E. Legge and Sheng He. I do not think I could have been where I am now without the support and guidance from Sheng, who believed in me from the beginning, encouraged me to reach my full potential and assured me through the difficulties of graduate school. I learned so much from him. I owe my deepest gratitude to Gordon, who helped me find my true passion in research, offered me endless support and patiently helped me find my way whenever I was lost, and pushed me beyond where I could have gone by myself. Gordon helped me grow immensely and I could not have asked for a better mentor.

I thank Drs. Daniel Kersten and Yang Zhang for serving on my committee and for their helpful feedback.

I also would like to extend many thanks to Mary Beth Kelly and the Learning Disabilities Association of Minnesota for their great help in participant recruitment, data collection, and administration and analysis of dyslexia assessments for the project in Chapter 4. I would not have been able to complete this project without their help and guidance.

A special thanks to Dr. Yingzi Xiong, who made an invaluable contribution to the project in Chapter 2, for her generous support, natural mentorship and wonderful friendship throughout my journey as a graduate student.

I would like to thank all my former and current lab-mates, specifically, Yueh-Hsun (Walter) Wu, Dr. Yingchen He, Dr. Aurélie Calabrèse, Rachel Gage, Christina Granquist, Andrew Freedman, Siyun Liu, Yijun Ge and Chen Chen. It is a great pleasure to know all of you; your companionship has not only been essential to my professional development but has also been crucial to have a fun and positive graduate school experience.

My sincere thanks to my dear friends, both in Turkey and the US for being by my side in every step of the way, supporting me throughout the most difficult times of my life and celebrating with me all the happy moments. One of the biggest thanks goes to Ummul Kathawalla, for truly shaping my life for the better in the US. I also would like to thank Yagizhan Yazar, Zara Wright and Mckenzie Kaubrys for their wonderful friendship over the last few years and making my Minneapolis experience unforgettable with so many amazing memories. My special thanks to Ezgi Cinar, whose presence and friendship made the last mile of the way so much more bearable and enjoyable. Last but not least, my deepest thanks go to Meltem Karaca, Bilge Barutcu and Oyku Kocoglu, who made the idea of time and space irrelevant, helped me remember my priorities in life and provided genuine and endless support, love and euphoria.

And I would like to give my special thanks to Scott Burwell. There are not enough words to express my gratitude for all the love, support and joy he brought to my life. He helped me grow academically, professionally and emotionally; and shared the happiest and saddest moments -and everything in between- of my life over the past five and a half years. He has been the fundamental piece of my happiness and success.

Finally, my biggest gratitude is for my beloved parents, Dogan and Meryem Atilgan, and my best friend, my sister, Zeynep Atilgan. I am beyond thankful for their unconditional and unwavering love and support. I owe them for everything that I achieved.

Dedication

To my beloved parents and sister

Abstract

Reading is one of the most important daily visual tasks. Majority of people become expert readers at a very young age and perform this complex task effortlessly and automatically throughout most of their lives. However, in this “effortless” task, multiple complex processes must be efficiently carried out in a very short amount of time. That is why, reading processes and their underlying mechanisms have been a great interest for researchers. For decades, researchers have been studying the constraints that prevent people from becoming “experts” in reading. These constraints may result from many different sources such as sensory limitations, cognitive disabilities, linguistic difficulties and so on. This thesis focuses on some of the sensory constraints that affect reading performance and present three studies that approach these constraints from different angles. Chapter 1 gives an overview of the thesis. Chapter 2 addresses the close relationship between text properties and reading performance. Specifically, in this chapter, the joint impact of print-size and display-size limitations on reading speed is investigated. The findings indicate that a minimum number of characters per line, which is determined by print-size and display-size constraints together, is required to achieve a criterion of 80% of maximum reading speed. Chapter 3 discusses the neural substrates of crowding, the inability to recognize objects when surrounded by other objects, which is directly related to reading performance given the nature of text materials (i.e. letters are surrounded by other letters). Specifically, the relationship between object recognition performance in crowding settings and the nature of visual parallel pathways (i.e. Magnocellular and Parvocellular pathways) were examined. The results showed that the magnitude of crowding is different depending on the visual stimuli and their relative engagement in the visual parallel pathways. For

example, processing of spatial forms like letters is affected by crowding the most while color processing does not seem to be impaired by the crowding effect. Finally, Chapter 4 investigates the visual limitations in reading through testing people with dyslexia. Visual span, the number of identifiable letters at one fixation, and visual crowding were studied to examine the differences between typical readers and individuals with dyslexia. In addition, various standardized assessments and Flashcard reading speed tests were administered to examine the individual differences, and potentially the sub-groups of dyslexia. Chapter 4 aims to address the relationship between individual differences and the performance in visual-attentional tasks in dyslexia. The findings indicated that despite large individual variabilities in the assessments and visual experiments, the dyslexia group on average demonstrated slower reading speeds, narrower visual span profiles and larger critical spacing values compared to the control group.

Together, these three studies provide a better understanding of sensory factors that limit reading performance.

TABLE OF CONTENTS

LIST OF TABLES	**VIII**
LIST OF FIGURES	**IX**
CHAPTER 1: OVERVIEW	**1**
CHAPTER 2: IMPACT OF PRINT SIZE AND DISPLAY SIZE ON READING	**9**
INTRODUCTION	**9**
MATERIALS AND METHODS	**14**
PARTICIPANTS	14
MATERIALS	15
PROCEDURE	19
ANALYSIS	20
RESULTS	**26**
READING WITH NORMAL VISION	26
READING WITH ARTIFICIALLY REDUCED ACUITY	28
READING WITH LOW VISION	30
DISCUSSION	**31**
CHAPTER 3: CROWDING AND PARALLEL VISUAL PATHWAYS	**37**
INTRODUCTION	**37**
EXPERIMENT I	**41**
METHODS	41
RESULTS AND DISCUSSION	45
EXPERIMENT II	**47**
METHODS	47
RESULTS AND DISCUSSION	53
GENERAL DISCUSSION	**54**
CHAPTER 4: CROWDING AND VISUAL SPAN IN DYSLEXIA	**59**
INTRODUCTION	**59**
METHODS	**65**
PARTICIPANTS	65
STIMULI, APPARATUS AND PROCEDURE	67
ASSESSMENTS	68
FLASHCARD READING SPEED TEST	72

Crowding Experiment 73
Visual Span Experiment 77
Data Analysis 79
Results **81**
Assessments 81
Flashcard Reading speed 84
Crowding 85
Reduced Visual Span Size in Dyslexia 87
Summary of Results 92
Discussion **93**
Crowding Effect in Dyslexia 93
Visual Span Profiles in Dyslexia 95
Future Directions 97

REFERENCES **98**

APPENDIX 1: INDIVIDUAL READING CURVES OF NORMALLY SIGHTED SUBJECTS IN CHAPTER 2 **106**

APPENDIX 2. ESTIMATING THE NUMBER OF CHARACTERS PER LINE IN CHAPTER 2 **108**

APPENDIX 3: RELATIONSHIP BETWEEN CRITICAL PRINT SIZE AND VISUAL ACUITY IN CHAPTER 2 **110**

APPENDIX 4: FONT PROPERTIES IN CHAPTER 2 **111**

APPENDIX 5. CHARACTER COUNTS AND PRINT SIZES IN CHAPTER 2 **112**

APPENDIX 6: THE INDIVIDUAL SCORES IN ASSESSMENTS AND READING SPEED TEST IN CHAPTER 4 **113**

APPENDIX 7: INDIVIDUAL CRITICAL SPACING VALUES IN CHAPTER 4 **114**

APPENDIX 8: CORRELATION COEFFICIENTS FOR ALL THE MEASUREMENTS IN CHAPTER 4 **115**

List of Tables

Table 2- 1 ______ *16*

Table 2- 2 ______ *27*

Table 2- 3 ______ *36*

Table 4- 1 ______ *67*

Table A4- 1 ______ *111*

Table A5- 1 ______ *112*

Table A6- 1 ______ *113*

Table A7- 1. ______ *114*

Table A8- 1 ______ *115*

List of Figures

Figure 2- 1 ______ *13*

Figure 2- 2 ______ *19*

Figure 2- 3 ______ *23*

Figure 2- 4 ______ *24*

Figure 2- 5 ______ *35*

Figure 3- 1 ______ *43*

Figure 3- 2 ______ *46*

Figure 3- 3 ______ *50*

Figure 3- 4 ______ *54*

Figure 4- 1 ______ *73*

Figure 4- 2 ______ *75*

Figure 4- 3 ______ *76*

Figure 4- 4 ______ *78*

Figure 4- 5 ______ *82*

Figure 4- 6 ______ *83*

Figure 4- 7 ______ *84*

Figure 4- 8 ______ *85*

Figure 4- 9 ______ *86*

Figure 4- 10 ______ *88*

Figure 4- 11 ______ *89*

Figure 4- 12. ______ *90*

Figure 4- 13 ______ *91*

Figure A1- 1 ______ *106*

Figure A1- 2 ______ *107*

Figure A2- 1 ______ *109*

Figure A3- 1 ______ *110*

Chapter 1: Overview

Reading is an important daily task, requiring coordination of multiple processes. There are many constraints that prevent individuals from becoming an expert in the complex reading task. These constraints may be grouped into two main categories: sensory and cognitive/linguistic.

Visual acuity is one of the main factors that is involved in sensory limitations for reading. Normally sighted people make use of their visual periphery to “pre-view” the upcoming letters and words. It is known that visual acuity is poor in peripheral vision. Therefore, poor visual acuity is relevant to both normally sighted and visually impaired individuals in reading. Another important sensory constraint that is relevant to both normally sighted and low vision people is crowding. Crowding is described as the inability to recognize objects that are in clutter (Levi, 2008). Crowding effect becomes especially strong in peripheral visual field and plays an important role in letter recognition in the periphery, and reading overall.

In addition to acuity limitations and crowding, external physical factors may also influence reading performance such as print size of the text, page or screen size that the text is presented on, contrast polarity, etc. Sensory limitations such as the impact of physical text properties, visual acuity limitations and constraints due to crowding, have been extensively studied in the reading literature (Legge, Mansfield, & Chung 2001; Pelli et al., 2007; Martelli et al., 2009). Although researchers are able to separately investigate the impact of these factors, the interactions between these constraints and their combined impact on reading are still unclear. Nevertheless, various psychophysics experimental designs are used as effective tools to improve our understanding of sensory limitations in reading.

On the other hand, cognitive/linguistic factors are mainly tested through standardized assessments, as it is challenging to manipulate the experimental parameters for such complex factors. Sentence context, word frequency, working memory capacity, phonological awareness and attentional abilities are some of the cognitive/linguistic factors that shape reading performance. Similar to sensory constraints, these cognitive factors frequently interact with each other. Limitations due to interactions between some of these factors give rise to specific reading difficulties such as dyslexia. Dyslexia is a specific reading disability characterized by poor word recognition and poor decoding skills, despite normal intelligence (American Psychiatric Association, 2013). Individuals with dyslexia are assumed to have both sensory and cognitive difficulties. It is important to understand how sensory and cognitive factors influence reading abilities and identify their mechanisms. This is the major motivation for this thesis research.

This thesis is composed of three separate studies formatted as journal articles. The following sections provide an overview of the three studies.

Study 1

One of the main determinators of reading performance is the properties of text. Whether it is displayed on a book page, a newspaper or a computer screen, the physical characteristics of the text influence individuals' reading performance. The size of the text has been widely studied and shown to greatly influence reading performance, which is mainly measured by reading speed in words per minute (wpm) unit. Previous studies demonstrated that reading speed starts to decline at a particular print size that is too small for maximum reading speed of the individual (Mansfield & Legge, 2007). This particular print size is defined as Critical Print Size (CPS). A few studies also investigated the effect of display size and found that reading speed is influenced by how big or small the display is (Fine & Peli, 1996; Beckmann & Legge, 1996). Display size and print size together limit the number of characters that can fit in a given line. These constraints become especially important in two situations: when people with normal vision read text on small digital displays, and when people with low vision read magnified text.

In chapter 2, we investigated how display size and print size interact and jointly affect reading performance. To do this, we measured reading speed as a function of print size in three different display formats, which are chosen based on commonly used daily devices (i.e. phone, tablet and laptop). For each display format, participants read sixteen different stories which were presented with a range of print sizes and in two different fonts. Reading performance was modelled to obtain critical values for both print size and display size limitations.

We found that reading performance is jointly determined by the print size and display size limitations for a given individual. Based on these findings, in Chapter 2, we present a unified framework for evaluating the joint impact of print size and display size constraints. This framework demonstrates that a minimum number of characters per line is required to achieve 80% of maximum reading speed. This number is defined as the critical character count (CCC). The CCC is 13 characters for normally sighted people and 8 for low vision people. This value is nearly consistent across font and display formats. CCC is likely to be due to oculomotor constraints, indicating the impact of perceptual span, which is defined as the region in which individuals obtain useful information at a single eye fixation in reading.

Together, our results show that for an individual to reach their maximum reading speed, the print size and the display size should be large enough to accommodate both CPS and CCC constraints.

Study 2:

Visual crowding, the inability to recognize objects in clutter, is one of the main sensory limitations for letter recognition (Whitney & Levi, 2011). Crowding is described as the bottleneck of object recognition due to its critical impact (Levi, 2008). Crowding is especially strong in peripheral vision. Reading makes great use of the information in the periphery, as individuals “pre-view” this information before bringing it to their central vision, therefore, crowding is highly relevant to reading performance. The impact of crowding on reading has been extensively studied (Pelli et al., 2007; Levi, Song, & Pelli, 2007; Chung, 2007; Martelli et al., 2009). Despite the extensive literature on crowding, the underlying neural mechanisms are still not well-understood. Previous studies mostly focused on investigating at what point in the visual hierarchy crowding takes place (Pelli, 2008; Anderson et al., 2012) and multiple studies have shown that crowding may occur at multiple stages of visual processing (Louie, Bressler & Whitney, 2007; Fischer & Whitney 2011; Anderson et al., 2012; Manassi & Whitney, 2018). Nonetheless, there is still much work needed to understand the neural correlates of crowding. Better understanding of these mechanisms will provide insight for the bottleneck of the object recognition and its limitations in reading.

In Chapter 3, we asked whether the visual crowding effect is different in visual parallel pathways. Given that visual parallel pathways, Magnocellular (M) and Parvocellular (P) pathways, represent biases toward different stimulus properties, and are involved in the processing of visual stimuli to different extents, the nature of the crowding effect might differ across these pathways. Understanding these differences, if there is any, would help understand the neural correlates of crowding better.

We tested the possible differences in crowding with regards to parallel visual pathways by designing the stimuli to selectively engage the P pathway or the M pathway. Chapter 3 presents two experiments; each uses different stimulus properties to separate processing in the two pathways. In the first experiment, high temporal frequency flickering and color are used to either to define the target or the background depending on the M or P pathway conditions. In the second experiment, color, motion and form recognition tasks are used to separately engage the two parallel pathways. In both experiments, subjects' overall performance and critical spacing values were evaluated in different conditions. Critical spacing is the minimum distance required between target object and nearby objects for complete recognition.

We found that at the same eccentricity and with the same type of visual tasks, objects processed in the M pathway appeared to be more vulnerable to crowding compared to the objects processed in the P pathway. Our results also showed that, when different types of stimuli and visual tasks (i.e. form, color and motion) are used, presumably involving different degrees of dependence on the M and P pathways, the strength of crowding is different depending on the stimulus categories. While color recognition performance does not seem to be affected from crowding, the most severe crowding effect is observed in the processing of spatial forms. These findings indicate that processing in the M pathway is more affected by crowding compared to the P pathway at a given eccentricity; and the processing of the form of the stimuli, which involves processing of letters, seems to be most affected by crowding compared to color and motion processing.

Study 3:

Dyslexia is a specific learning disorder, characterized by difficulties in reading. Individuals with dyslexia show impairments in identifying speech sounds, relating letters to their corresponding sounds, visuo-spatial attentional skills and working memory capacity. Dyslexia is defined as a multi-syndrome disability, as individuals with dyslexia greatly differ in their symptoms and in the severity of their difficulties. Nonetheless, there are two main impairments that can be observed in the majority of individuals with dyslexia: poor phonological representations and difficulties in visual attentional tasks. The latter has been a great interest for vision researchers and it also motivated the last chapter of the current thesis.

In Chapter 4, we investigated the extent of crowding and properties of visual span profiles of both typical readers and individuals with dyslexia. Visual span is the number of letters that can be recognized without eye movements and it is demonstrated to be an important factor affecting reading performance (Legge et al., 2001; Legge et al., 2007). Previous studies showed evidence for a high correlation between visual span size and reading speed (Yu, Cheung, Legge & Chung, 2007). Surprisingly, to our knowledge, the relationship between dyslexia and visual span has not been investigated prior to this thesis.

In addition to the visual span and crowding experiments, we also investigated the reading and phonological abilities of all participants through standardized tests. A wide battery of assessments from Woodcock Johnson-IV (WJ-IV) Tests of Achievement, Test of Auditory Processing Skills 3 (TAPS-3) and Castle and Coltheart Test 2 (C&C2) were administered. Participants' reading speeds were also measured in a Flashcard reading test.

We found that individuals with dyslexia on average scored lower in all of the assessments, although individuals in both dyslexia and control groups showed large variabilities within their groups. Individuals with dyslexia were found to have smaller visual span size on average compared to typical readers. Lastly, the critical spacing values, which were obtained in crowding experiments similar to those in Chapter 3, were larger in the dyslexia group compared to the control group, indicating that individuals with dyslexia required larger spacing between the target object and nearby objects for recognition. This finding suggests that individuals with dyslexia suffer from stronger crowding effects.

Summary

The studies presented in this thesis addressed the wide range of sensory limitations that impact reading performance. Three major conclusions emerge from this thesis:

1- For a particular individual, either with normal or low vision, a minimum print size (CPS) and a minimum character count (CCC) are required to reach maximum reading speed. Under the circumstances that these requirements are not met such as large magnification for low vision or small digital displays for normal vision, reading speed is slower than maximum reading speed.

2- Crowding, defined as the bottleneck of object recognition and shown to impact reading performance, shows differences in regard to parallel visual pathways. Processing of spatial forms, which is assumed to engage both P and M pathways, is affected by crowding the most compared to the color (engaging P pathway) and motion (engaging M pathway) processing. With the same visual tasks and at a given eccentricity, the M pathway seems to be more vulnerable to crowding compared to the P pathway.

3- Individuals with dyslexia show large variability in reading and phonology assessments in terms of both characteristics and severity of the impairment, suggesting that dyslexia is a spectrum in reading difficulties rather than an unified single condition. On average, individuals in the dyslexia group have larger critical spacing values for letter and symbol stimuli, suggesting stronger crowding effects in dyslexia compared to typical readers. In addition, the visual span profiles in the dyslexia groups were narrower compared to the control group.

These results provide insights into the sensory constraints on reading performance and help our understanding regarding the underlying neural mechanisms of these limitations.

Chapter 2: Impact of Print Size and Display Size on Reading

Introduction

Two fundamental constraints limit the number of characters that can be displayed in text at one time--print size and display format. The print size must be legible for the reader, and the size of the display (or page) limits the amount of text that can be rendered at this print size. As the print size gets larger, the amount of displayable text (number of characters per line and number of lines per page or screen) shrinks. These dual constraints conflict in two important situations—when magnification is required for people with low vision, and when people with normal vision read text on small digital displays. In this paper, we provide a unified analysis of the joint impact of these constraints. We also present empirical evidence showing how these constraints limit reading performance in cases of reduced acuity and small displays.

A widely used measure of text legibility is reading speed, measured in words per minute (Tinker, 1964; Carver, 1976; Legge et al., 1985). Reading speed is straightforward to measure, is sensitive to changes in both eye condition and text properties, and is functionally significant to readers (Legge, 2007). The relationship between print size and reading speed has been studied in detail, reviewed by Legge and Bigelow (Legge & Bigelow, 2011). Numerous studies have shown that as angular print size (i.e. the visual angle subtended by text letters) increases from the reader's acuity limit, reading speed increases until a critical print size (CPS) is reached, and then levels off at a maximum reading speed (MRS) for print sizes larger than the CPS. An example of reading speed as a function of print size is shown in Figure 1a. This typical reading speed curve has been verified by various studies and the idea of CPS is widely used by researchers and clinicians. For normally sighted readers, the CPS is approximately 0.2º, and reading speed remains maximum for a factor of ten in print size from 0.2º to 2º (Legge & Bigelow, 2011).

Early studies by Tinker and Paterson showed that when print size is fixed, the length of text lines, measured in picas (one pica = 0.167 inch), affects reading speed, indicating physical line length is an important factor to be considered when deciding on typographical layout (Tinker & Paterson, 1931). Several later studies examined the effect on reading speed of "window size" or "field of view" of magnifiers (Fine & Peli, 1996; Beckmann & Legge, 1996), and provided estimates of the minimum field size in terms of the number of characters visible in the magnifier's field of view. These prior findings are suggestive of the impact of display format on reading speed, but do not show how print size and display size jointly constrain reading performance for continuous text. In the current study, we first examined the hypothesis that for an individual to achieve maximum reading speed, lines of text must include at least a critical number of characters. We term this hypothetical number the critical character count (CCC). Our hypothetical curve of reading speed as a function of character count per line is shown in Figure 1b: the reading speed stays at its maximum for large character counts but drops for character counts below the CCC. We hypothesize that the CCC (green vertical line) determines the minimum size of displays for effective reading.

Why would the number of characters per line affect reading speed? Some property of the text, unrelated to the reader's vision status, might impose a constraint. For example, the distribution of word lengths might be crucial; reading speed may be unaffected as long as the line length can accommodate most or all of the words in the text, but may slow down when some individual words occupy more than one line. If an intrinsic text property is the limiting factor, we might expect similar CCC values for participants with both normal and low vision. Alternatively, the character count impact on reading speed might be determined by the perceptual span. McConkie and Rayner defined the perceptual span as the region around fixation in which printed information facilitates reading behavior (McConkie & Rayner, 1975). They introduced a “moving window” method to measure the perceptual span in which gaze-contingent eye tracking was used to distort text at varying distances from the point of fixation. Studies have shown that the perceptual span in normal vision includes three or four characters to the left of fixation and 14 to 15 characters to the right of fixation (Rayner & Bertera, 1979; Rayner, 1986). It is measured in terms of character spaces because it is independent of angular print size over a wide range (Morrison & Rayner, 1981; O’Regan, 1983; Miellet et al., 2009) and is not font dependent (Rayner et al., 2010). Recent studies have shown that the perceptual spans of low-vision participants with macular degeneration are substantially smaller than in normal vision (Crossland & Rubin, 2006; Calabrese et al., 2014; Calabrese et al., 2016). It is plausible that if lines of text have fewer characters than the extent of a reader’s perceptual span, reading would slow down because less information is available on each eye fixation. Moreover, if the size of the perceptual span determines the critical character count, we would expect the CCC to be lower in low vision than in normal vision.

Interest in the text capacity of small displays emerged with the advent of digital displays on microwave ovens and other appliances, and then with mobile devices such as cellphones and smart watches. Similar concerns exist with traffic displays and other electronic message signs viewed at a distance. For a given print size, the screen size determines the number of characters per line and the number of lines on the display, and therefore the total text capacity for the display. When the display capacity is small, many pages will be required to render lengthy texts, with associated time costs in line and page switching.

The tradeoff between print size and screen size becomes particularly acute for people with low vision who require large print to read. By a recent estimate, there are 5.7 million Americans with impaired vision, with the number expected to increase to 9.6 million by the year 2050 as the population ages (Chan et al., 2018). Most people with impaired vision are not blind, but have low vision. They continue to read visually, but require substantial magnification of print. There is an important need for enhancing accessibility of websites and other digital displays for low-vision users by providing customizable text formats in terms of the number of characters per line and lines per screen. The flexibility of digital displays for customizing print size, page layout and other properties of text has substantial advantages for people with low vision (Legge, 2016; Wu et al., 2020). But digital displays on small, mobile devices pose challenges for people with low vision. For example, suppose a small display can fit 10 lines of 60 characters per line at the CPS of a normally sighted reader. The same display might only accommodate 1 line of 6 characters for a person with 20/200 acuity.

The major goal of the research presented in this paper was to establish how display format interacts with the need for adequate print size in constraining reading performance for people with both normal and low vision. Our hypothesized a unified framework for evaluating the joint impact of these constraints, shown in Figure 2-1c & d. Critically, for a given display format and font, the angular print size (lower axis) determines the character count per line (top axis); as the print size increases, the character count decreases. This reciprocal relationship enables the independent constraints on reading speed of print size (red curve from Figure 2-1a) and character count (green curve from Figure 2-1b) to be represented in a unified framework (Figure 1c & d).

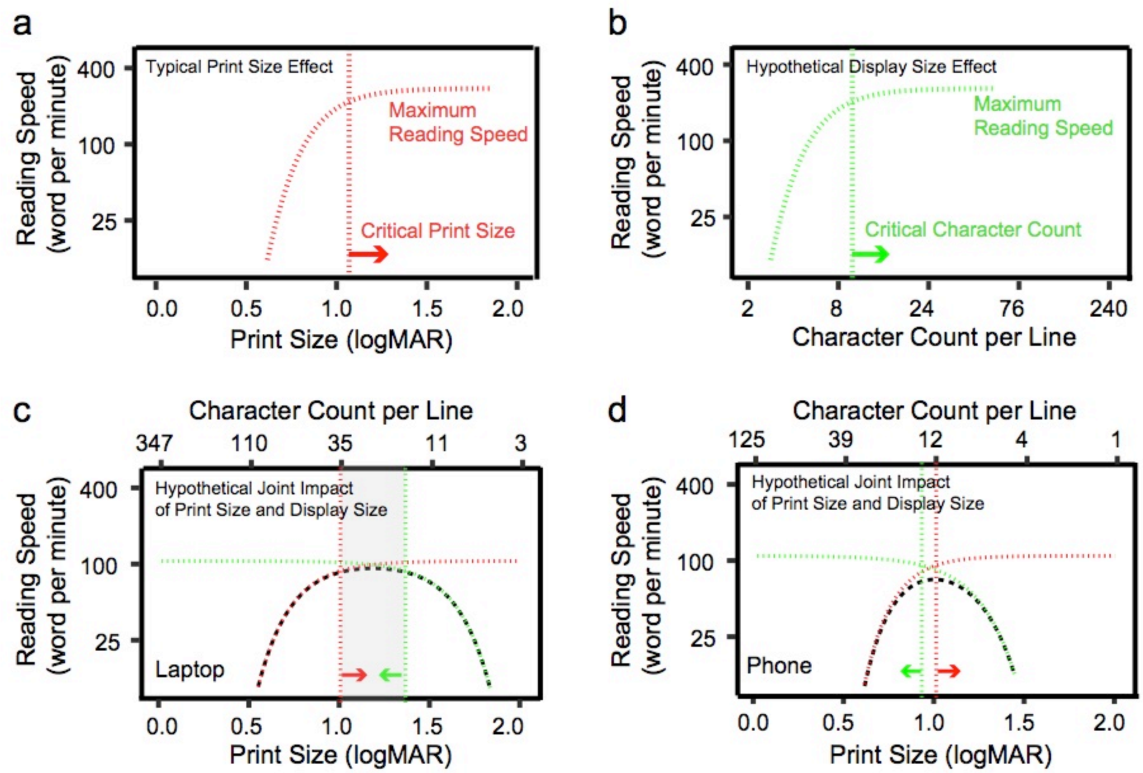


Figure 2- 1: Illustration of the impact of print size and display size on reading speed. a. A typical reading curve illustrating the impact of character print size on reading speed (3,4,22). b. A hypothetical reading curve illustrating the impact of display format on reading speed. c & d. The hypothesized reading curves showing the joint impact of print size and display format (c: laptop, d: phone) on reading speed. The number of characters per line is now expressed in terms of angular print size. The conversion is described in Appendix 2. The CPS and the print size corresponding to the CCC determine the lower and upper bounds of the range of recommended print size that allows near-maximum reading speed. In the examples shown in c, a recommended print size range exists in the laptop format, as indicated by the grey zone. However, as shown in d, near-maximum reading is not possible for the phone format because the CPS exceeds the CCC.

The black curves show the impact of the joint constraints, indicating that reading speed is expected to fall on or below the red and green curves. The CPS (red vertical line) determines the smallest print, the CCC (green vertical line) determines the largest print, and the gray zone in between represents the range of print sizes for achieving near-maximum reading speed (Figure 2-1c). When a large CPS is required on a small display, the gray zone will shrink, and entirely disappear if the CPS exceeds the print size associated with the critical character count (Figure 2-1d). In this case, readers cannot achieve their maximum reading speed.

To examine this unified framework, we measured reading speed as a function of print size for participants with both normal and low vision. They were tested with eight print sizes and three display configurations simulating typical sizes for cellphones, tablets and laptops. The eight print sizes were selected to approximately match the character counts across the three display formats (see Figure 2-2 and Methods). Participants were instructed to read silently, as quickly and accurately as possible, while retaining good comprehension. Reading speed was calculated as the total number of words read within a one-minute time period. We compared how the joint impact of print size and character count on reading speed changes with vision status, display format, and font.

Materials and Methods

Participants

Thirty normally sighted students (mean age = 21.6 years) participated through the University of Minnesota Research Experience Program. They had normal or corrected to normal vision, with no history of reading impairments. The participants were separated into two groups (N = 14 and N = 16) to read with different fonts (see below). The sample size was determined based on our primary interest in the impact of display formats on the critical number of characters per line for near-maximum reading. From previously published data, we estimated the distribution of the number of characters per line at the preferred print size and viewing distance for a group of normally sighted participants (Granquist et al., 2018). Using the "SIMR" package, we obtained the minimum sample size yielding a significant difference in display format at 80% power, based on 1000 simulations of each sample size (Green & MacLeod, 2016; Brysbaert, & Stevens, 2018). The result showed that to achieve 80% power at $p = 0.05$, with an effect size of Cohen's $d = 1.15$, a sample size of at least 10 is needed.

Ten low vision (mean age = 58.3 years) participants were recruited from the Minnesota Laboratory for Low-vision Research roster. These participants had heterogeneous diagnoses and levels of visual impairment. They were selected because of their ability and interest in reading large print regardless of diagnosis. The diagnosis and binocular distance visual acuity for the low-vision participants are provided in Table 1. They were all native-English speakers and had no history of dyslexia or other reading disabilities. All participants gave written informed consent. The study was approved by the University of Minnesota Institutional Review Board (IRB).

Materials

<u>Stimuli</u>

Forty-eight short stories from Grimms' Fairy Tales (The Brothers Grimm, 2008) were used in this study. The stories were screened to avoid offensive content and were selected from the full set of stories in the book. The forty-eight stories have similar levels of complexity in vocabulary, content and style. The average Flesch-Kincaid grade score of the stories is seventh grade (calculated with text readability consensus calculator).

The stories were rendered in either the Times New Roman or Courier font. Normally sighted participants were randomly assigned to one of the two font conditions, resulting in 14 participants in the Times New Roman and 16 in the Courier conditions. All the low vision participants were tested with the Times New Roman font.

ID	Diagnosis	Visual Acuity	Display	MRS wpm	Constrained-MRS wpm	CPS logMAR	CCC	CCC logMAR	Range logMAR
01	Left eye glaucoma, no sight right eye	0.48	laptop	182	182	0.49	11	1.64	1.15
			tablet	204	204	0.49	8	1.61	1.12
			phone	240	234	0.47	11	1.20	0.73
02	Retinal Detachment	0.54	laptop	263	257	0.80	11	1.65	0.85
			tablet	269	269	0.82	8	1.62	0.80
			phone	275	257	0.74	11	1.22	0.48
03	Septo-optic Dysplasia	0.68	laptop	102	100	0.90	9	1.75	0.84
			tablet	95	95	0.90	7	1.68	0.78
			phone	105	95	0.87	8	1.32	0.45
04	Glaucoma	0.68	laptop	32	32	0.65	8	1.79	1.14
			tablet	31	31	0.65	7	1.70	1.05
			phone	30	30	0.62	7	1.38	0.75
05	Macular Hole	0.74	laptop	355	331	1.07	11	1.65	0.58
			tablet	355	339	1.08	8	1.62	0.54
			phone	398	316	1.01	11	1.22	0.20
06	Diabetic Retinopathy	0.84	laptop	178	170	1.06	8	1.77	0.71
			tablet	162	158	1.07	7	1.69	0.62
			phone	162	145	1.01	8	1.35	0.34
07	Aniridia	1.08	laptop	204	186	1.33	8	1.77	0.44
			tablet	195	182	1.31	7	1.69	0.37
			phone	209	138	1.32	8	1.34	0.02
08	Aniridia	1.12	laptop	117	115	1.16	8	1.76	0.61
			tablet	117	112	1.14	7	1.68	0.54
			phone	117	95	1.14	8	1.34	0.20
09	Age-related Macular Degeneration	1.14	laptop	123	74	1.84	8	1.78	-0.06
			tablet	115	72	1.74	7	1.69	-0.05
			phone	120	44	1.63	8	1.36	-0.28
10	Diabetic Retinopathy, Glaucoma	1.16	laptop	148	129	1.39	9	1.76	0.37
			tablet	145	126	1.38	7	1.68	0.30
			phone	141	76	1.44	8	1.34	-0.10

Table 2- 1: The individual diagnosis, visual acuity and reading indices of the low vision participants. The reading indices include maximum reading speed (MRS), constrained-maximum reading speed (constrained-MRS), critical print size (CPS), critical character count (CCC) and the range of recommended print size (Range).

Participants read different stories in twenty-four conditions defined by eight print sizes and three display formats, as described below. While all the participants read two stories for each condition, normally sighted participants were tested with artificial acuity reduction (blur) for the second set of twenty-four stories. Blur was produced with customized diffusing goggles. Specifically, three layers of polyethylene films were added in front of a pair of safety goggles. The diffusing film layers were carefully flattened and tightly stretched to avoid optical distortion. These goggles artificially reduce acuity to an average of 0.83 logMAR (Snellen 20/135), measured with the Lighthouse Distance Visual Acuity chart, and reduce Pelli-Robson contrast sensitivity to an average of 1.0 log unit. LogMAR (Logarithm of the Minimum Angle of Resolution) is a unit of angular print size, which refers to the retinal-image size and depends on both the physical size of the print and the subject's viewing distance. The LogMAR value can be calculated from the angular print size (in degrees) by the following equation: LogMAR = $\log_{10}$(angular print size/0.083).

No story was presented in more than one condition for any subject. The pairings of story and presentation condition were randomly selected for each subject. The order of the conditions was counterbalanced across participants.

<u>Display</u>

All the stories were presented on a 27-inch Apple Cinema Display (2560×1440 pixels, pixel density:109ppi, refresh rate:60Hz). The stories were displayed with black text (0.42cd/m2) on a white background (432cd/m2).

In separate conditions, text was confined to portions of the monitor simulating the size of commonly used digital displays—a laptop (6.6 x 11.4 inches, matching a 13-inch MacBook Pro), a tablet (5.9 x 7.9 inches, matching an iPad), and a cellphone (2.3 x 4.1 inches, matching a iPhone 6).

Each story was presented in one of the eight print sizes and one of the three display formats. For a given print size, the number of characters per display (the display character count) varied with the size of the display. We selected a different set of eight print sizes for each display format. For each format, the smallest print size was 12 pt. The remaining print sizes were chosen to approximately match the sets of character counts. All three display formats were presented in landscape layout. This layout allowed us to minimize the possible effect of word splitting. Word splitting was only observed in the largest print size condition in cellphone and tablet displays. For words longer than 6 characters (approximately 10% of the words in our reading material), the first 5-6 characters were displayed in one page, and the following characters appeared in the next page, no hyphens were used.

Figure 2-2 shows two sets of sample stimuli, one set demonstrates stories with equal character counts per page across the three display formats, while the other set demonstrates stories with similar print sizes across the display formats. The specific ranges for print sizes (in units of pt and logMAR) and character counts per page for each display and font are provided in Appendix 5.

Figure 2- 2: Two sets of sample stimuli. The upper panel shows a story excerpt with equal character count per page across three display formats. The lower panel shows excerpts with similar print size across the three display formats. Display formats and print sizes in the figure are scaled in size to fit journal requirements. Here, the same story was used across the displays for demonstration purposes. In the actual experiment, no story was presented more than once to a given subject.

Procedure

Participants were tested with the Lighthouse near letter acuity chart before the main experiment. Participants were seated at a specified viewing distance from the display (normally sighted participants at 60cm, low-vision participants at 40cm).

At the beginning of the study, participants were told that their comprehension would be evaluated, and they were instructed to silently read the story presented on the screen for one minute, as fast and accurately as possible, switching pages when necessary. They pressed the spacebar to change the page. There was no noticeable delay associated with page switching. After one minute, a sound indicated the end of the trial and the subject reported the last word read. The experimenter recorded the ending word. This information was used later to calculate the reading speed (see Analysis section). For several randomly selected stories, participants were asked comprehension questions to confirm they were understanding the stories. Each subject was asked fifteen comprehension questions across 48 stories. No subject made more than 3 mistakes in the comprehension questions.

Prior to the main experiment, each subject was tested with a practice trial taken from "The Catcher in the Rye", presented in Times New Roman in the tablet display format with the 4th largest print size of the display (1.24 logMAR, 35 character per page). The comprehension test was not included in the practice trial.

After the practice trial, participants started the main experiment with one of the three display formats. Each display format included eight print sizes and within each display format, the stories were presented either from largest to smallest print size or vice versa. This pattern was repeated six times (two times for each of the three display formats). Normally sighted participants either started with the blur goggles on and took them off after the first three display conditions (total of twenty-four stories) or started without the blur goggles and put them on after the first half. All the participants were encouraged to take breaks between the stories.

Analysis

<u>Reading Speed</u>

Reading speed (words per minute, wpm) was calculated based on the last word read by participants when the one-minute time limit ended. Specifically, the experimenter calculated how many characters (including spaces) were read in each 1-minute trial using the character count feature in Microsoft Word. The character count was divided by six to estimate the number of “standard-length words” read (Carver, 1976). Reading speeds were obtained at each of the eight print sizes tested in each condition (i.e., three display formats x two blur/no-blur conditions for normally sighted participants; and three display formats for low-vision participants)

Reading Model

We modeled the joint impact of print size and display format on reading speed by incorporating the existing model for print size and our hypothetical model for character count.

It is well known from the reading literature that reading speed drops with print size in an exponential manner (Legge et al., 1985; Mansfield et al., 1996; Legge, 2007). Specifically, the reading speed stays at its maximum (MRS) for large print sizes but drops rapidly for small print sizes (Figure 2-1a). We implemented the exponential function in our model to describe the independent impact of print size on reading speed in cases where display size is sufficiently large so that its impact is negligible (Eq.1).

$$rs = MRS\left(1 - e^{-e^{lrc_{ps}(PS - xint_{ps})}}\right); \qquad \text{(Eq.1)}$$

where rs is the reading speed in log wpm (word per minute), lrc_{ps} is the rate of change in reading speed with print size (in logMAR), and $xint_{ps}$ is the print size at which reading speed is 0 log unit.

The impact of character count on reading speed, however, is not well established. We hypothesized that character count affects reading speed in a similar way as print size: the reading speed stays at its maximum for large character counts but drops for small character counts. We used a similar exponential function to describe the independent impact of character count on reading speed, in cases where the print size is sufficiently large so that its impact is negligible (Eq.2).

$$rs = MRS(1 - e^{-e^{lrc_{cc}(CC - xint_{cc})}}); \qquad \text{(Eq.2)}$$

where again rs is the reading speed in log wpm, lrc_{cc} is the slope of the rate of change in reading speed with character count per line, and $xint_{cc}$ is the character count at which log reading speed is 0 log unit. We verified the hypothetical pattern of Eq.2 by our empirical reading speed vs. print size data under the normal viewing condition (Figure 3a).

In real-life situations, both constraints are present. We hypothesized that trade-off between print size and the number of characters per line connects the two independent constraints and imposes a joint impact on reading speed. The hypothetical joint impact of print size and character per line on reading speed is modelled by Eq.3.

$$rs = MRS(1 - e^{-e^{lrc_{ps}(PS - xint_{ps})}} - e^{-e^{lrc_{cc}(CC - xint_{cc})}});$$

(Eq.3)

To express these constraints in terms of one independent variable, the number of characters per line is expressed in terms of angular print size, given specification of the display format and font. Briefly, when angular print size increases, the number of characters per line decreases; for a fixed angular print size, the number of characters per line is larger on a wider display and with a narrower font. A detailed derivation of the transformation from mean character count per line to print size in logMAR units is provided in Appendix 2. Eq.3 was used to fit the reading speed vs. print size data under the artificially reduced acuity (Figure 2-3b) and low vision (Figure 2-4) conditions.

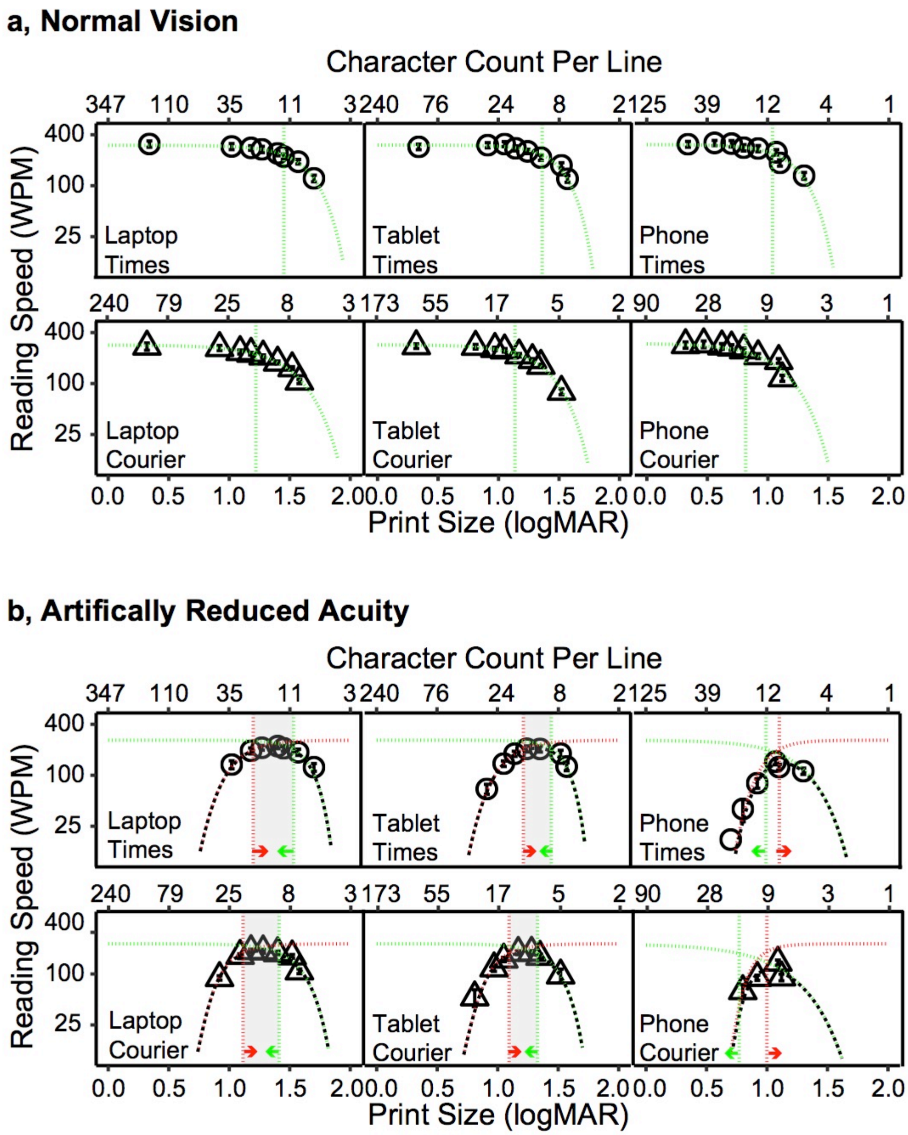


Figure 2- 3: The group average reading curves of the normally sighted participants in the normal viewing (a) and artificially reduced acuity (b) conditions. The reading speeds are represented by circles for Times and by triangles for Courier. The red and green dashed curves (fitted by Eq.1 and Eq.2) illustrate the impact of print size and character count on reading speed, respectively. The black dashed curve is the actual reading curve jointly affected by print size and character count (fitted by Eq.3). The red and green vertical lines represent the CPS and CCC corresponding to 80% of the maximum reading speed, and the gray area between them is range of print size for near-maximum reading speed (at 80% of the MRS), which we term the recommended print size range. In the normal viewing condition (a) all print sizes were larger than the CPS for normally-sighted participants (5), therefore only the green curves are plotted representing the character count effect. Notice that in some situations (e.g., reading with the phone format in the artificially reduced acuity condition), a recommended print size range doesn't exist, that is, there is no range of print sizes for which the subject can achieve at least 80% of MRS.

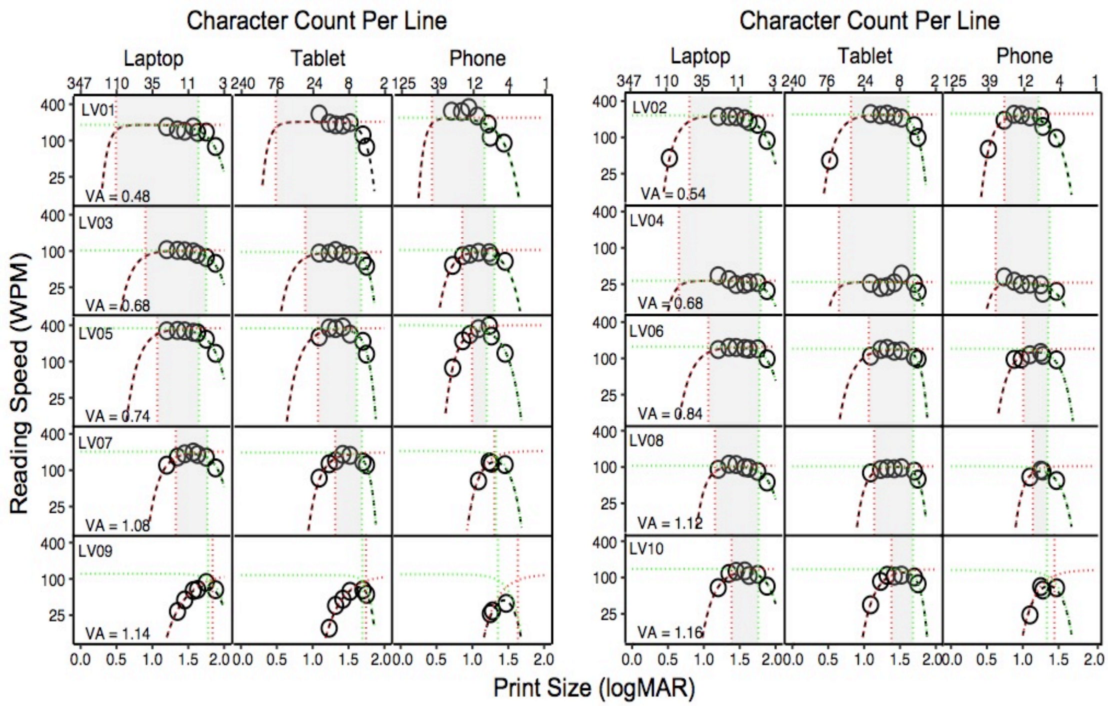


Figure 2- 4: The reading curves of low-vision participants. The measured reading speeds are represented by circles. The red and green dashed curves (fitted by Eq.1 and Eq.2) illustrate the impact of print size and character count on reading speed, respectively. The black dashed curve is the actual reading curve jointly affected by print size and character count (fitted by Eq.3). The red and green vertical lines represent the CPS and CCC corresponding to 80% of the MRS, and the gray area between them is the range of print size for near-maximum reading speed (at 80% of the MRS). Notice that for some combinations of severe vision loss and display format (e.g., all three displays for subject 9, and phone format for subject 10), there is no recommended print size range.

Curve Fitting and Key Reading Indices

Plots of reading speed vs angular print size were fitted using a Nonlinear Mixed Effects Model (NLME, “nlme” package) as described in Cheung et al (2008). The NLME model treated MRS, lrc_{ps}, lrc_{cc}, $xint_{ps}$, $xint_{cc}$ as fixed effects and subject as the random effect. Display was included as a covariate for the fixed effects and nested within subject in the random effect. Additionally, Font was included as a covariate for the normally sighted participants, and visual acuity was included as a covariate for the low vision participants. The significance of the covariates in the fixed effects were examined by the F statistics (ANOVA) in the nlme package, and the significance of the random effects was examined by the likelihood test, also using the ANOVA function. The non-significant components were excluded from the NLME model step-wise.

A display should provide a range of print sizes enabling maximum or near maximum reading speed. We adopted a criterion of 80% of maximum reading speed as “near-maximum reading speed”. Five key reading indices were obtained from the optimally fitting model, as summarized below:

- Maximum reading speed (MRS): the fastest reading speed participants can achieve without any constraint of print size or display format.
- Constrained maximum reading speed (constrained-MRS): the fastest reading speed participants can actually achieve when constrained by print size and display format.
- Critical print size (CPS): the smallest print size for a reading speed of 80% of the maximum reading speed.
- Critical character count per line (CCC): the smallest character count per line for reading speed of 80% of the maximum reading speed.
- Range of recommended print sizes: print sizes between the CPS and the print size associated with the CCC.

Linear mixed effect modeling (LME) was then performed to clarify the contributions of font and display format on each of the reading indices, and significant main effects and interactions were followed by pairwise comparisons with Bonferroni adjustment. Notice that the reading indices in the no blur and artificial blur conditions were analyzed in a single LME model to reduce Type-I errors.

Results

Reading with Normal Vision

We examined the impact of the number of characters per line on reading speed and evaluated the existence of a critical character count (CCC), in normal vision. Thirty normally sighted participants read different stories from *Grimms'* Fairy Tales in twenty-four conditions defined by eight print sizes and three display formats. Fourteen of the participants read with the Times New Roman font and sixteen read with Courier. The group reading curves for each combination of display format and font are shown in the panels of Figure 2-3a. Individual reading curves are provided in Appendix 1, Figure S1&S2.

The reading curves isolated the impact of character count on reading speed, because all of the tested print sizes were larger than the CPS for normally-sighted participants (Legge & Bigelow, 2011). As shown in Figure 2-3a and Figure S1, the reading speed remained constant for large character counts and dropped at smaller character counts, following the expected pattern. We modelled the reading curves as a function of character count (Eq.2, see Method). Two reading indices were obtained from the fitted curves: maximum reading speed (MRS) and critical character count per line (CCC), defined as the smallest character count per line for near-maximum reading speed of 80% of the MRS. The estimated values are provided in Table 2-2.

	Display, Font	MRS, wpm	Constrained-MRS wpm	CPS logMAR	CCC	CCC (logMAR)	Range* (logMAR)
No blur	Lap, T	300 [261,339]	--	--	12.8 [10.5,15.0]	1.42 [1.37,1.47]	--
	Tab, T	301 [264,339]	--	--	10.7 [8.5,13.0]	1.34 [1.30,1.37]	--
	Pho, T	306 [267,345]	--	--	11.5 [9.3,13.7]	1.02 [0.97,1.07]	--
	Lap, C	285 [247,323]	--	--	15.4 [13.3,17.5]	1.21 [1.16,1.27]	--
	Tab, C	287 [252,321]	--	--	12.3 [10.2,14.4]	1.14 [1.09,1.19]	--
	Pho, C	298 [253,342]	--	--	13.6 [11.5,15.7]	0.81 [0.76,0.87]	--
Blur	Lap, T	260 [232,287]	221 [194,249]	1.22 [1.18,1.25]	10.1 [7.8,12.3]	1.52 [1.49,1.55]	0.30 [0.25,0.35]
	Tab, T	260 [232,287]	210 [184,237]	1.22 [1.19,1.26]	8.5 [6.3,10.7]	1.43 [1.40,1.46]	0.21 [0.16,0.26]
	Pho, T	260 [232,287]	142 [120,163]	1.11 [1.07,1.14]	13.9 [11.7,16.2]	0.95 [0.89,1.02]	-0.15 †[-0.23,-0.08]
	Lap, C	227 [213,241]	188 [175,202]	1.12 [1.08,1.15]	9.8 [7.7,11.9]	1.39 [1.35,1.43]	0.27 [0.22,0.33]
	Tab, C	227 [213,241]	179 [162,195]	1.11 [1.06,1.16]	8.4 [6.3,10.5]	1.30 [1.25,1.34]	0.19 [0.12,0.25]
	Pho, C	227 [213,241]	114 [99,130]	0.99 [0.95,1.02]	18.9 [16.8,21]	0.71 [0.62,0.80]	-0.28 [-0.38,-0.17]
Low vision	Lap, T	170 [114,226]	157 [102,212]	1.07 [0.83,1.31]	9.1 [8.3,9.9]	1.73 ‡ [1.70,1.77]	0.66 [0.44,0.89]
	Tab, T	169 [111,227]	159 [102,217]	1.06 [0.83,1.29]	7.3 [7.0,7.6]	1.67 [1.64,1.69]	0.61 [0.39,0.83]
	Pho, T	180 [115,245]	143 [83,203]	1.03 [0.80,1.26]	8.8 [7.8,9.8]	1.31 [1.26,1.35]	0.28 [0.07,0.49]

Table 2- 2: Summary of the reading indices in the normal vision and low vision groups (Mean [95% Confidence Interval]). The reading indices include: maximum reading speed (MRS), constrained maximum reading speed (constrained-MRS), critical print size (CPS), critical character count (CCC) and its corresponding logMAR value, and the range of recommended print size (Range).

Our primary question was whether the CCC changes with font and display format. Font (Times or Courier) did not significantly affect the CCC ($F(1,28) = 3.74$, $p = 0.06$). The only significant display difference was between tablet and laptop when reading with Courier (Mean difference (MD) = 3.1, p= 0.005, 95% confidence interval (CI) = [0.7, 5.4]). The average CCC per line across fonts and displays was 12.8 characters.

The MRS ranged from 300 to 306 wpm for Times, and ranged from 285 to 298 wpm for Courier. Although this difference was not significant ($F(1,28) = 1.26$, $p = 0.27$), the slightly faster MRS in Times for normally sighted participants is consistent with previous findings (Mansfield et al., 1996).

Reading with Artificially Reduced Acuity

We next examined the joint impact of print size and character count on reading speed. We asked the same normally sighted participants to read another twenty-four stories while wearing goggles covered with diffusing films, which artificially reduced their acuity to an equivalent letter acuity of 0.83 logMAR (Snellen equivalent 20/135). The group reading curves are shown in Figure 3b. Individual reading curves are provided in Appendix 1, Figure S1&S2.

Reduced acuity necessitated larger print size for maximum reading, requiring attention to both print size and display format constraints. These joint constraints can be observed in Figure 3b, consistent with our expected pattern (Figure 2-1c&d). Accordingly, the reading speed was modelled as a function of both print size and character count (Eq.3, see method).

The mean MRS values were 260 and 227 wpm for Times and Courier, slower than the no blur condition ($F(1,140) = 42.3$, $p < 0.001$). The CCC was similar for laptop and tablet, averaging 9.2 characters, which was significantly smaller than the no-blur condition (laptop: $p < 0.001$, MD = 4.1, 95% CI = [-6.1, -2.2]; tablet: $p = 0.003$, MD = 3.0, 95% CI = [-5.0, -1.1]). The CCC for phone averaged 16.4 characters, which was significantly larger than the no-blur condition ($p < 0.001$, MD = 3.8, 95% CI = [1.9, 5.8]). The larger CCC of the phone display might have been an artifact of the simulation as we did not observe this effect with low vision participants, as shown in the low vision results below. In our joint model we considered CCC and CPS as two independent constraints (Eq.3). Consistent with this hypothesis, there was no significant correlation between CCC and CPS for any font or display (all $p > 0.05$).

There exists a range of print sizes that allows the participants to read at a reading speed of at least 80% of their maximum reading speed. This range is shown in gray in Figure 2-3b. The range averages 0.29 logMAR for laptop (nearly a factor of two in print size) and 0.20 logMAR for tablet (about a factor of 1.6 in print size). For our testing conditions, there is no print size enabling 80% of maximum reading speed for the cellphone display because the print size required to achieve the CCC is smaller than the CPS.

We refer to the highest reading speed achievable, given print size and display format constraints, as the constrained maximum reading speed. Here the mean constrained-MRS values were significantly lower than the MRS for all three displays. For laptop and tablet formats, the constrained-MRS was slower than MRS by 15.9% ($p < 0.001$, MD = 38.5 wpm, 95% CI = [32.6, 44.4]) and 20.1% ($p < 0.001$, 95%, MD = 48.7 wpm, CI = [44.7, 56.8]), respectively, indicating that although reading speed exceeded the criterion value (80% or more of the MRS), the joint constraints of print size and character count prevented participants from reaching their MRS.

For the phone format, the constrained-MRS was slower by 47.5% ($p < 0.001$, MD = 115.2 wpm, 95% CI = [103.8, 126.6]).

Reading with Low Vision

How do the joint constraints of print size and character count affect the reading speed of people with low vision? Ten low vision participants with various vision diagnosis and binocular visual acuities (listed in Table 2-1) participated in our study. These participants were chosen because they continue to read visually on a regular basis in their daily lives. They performed the story reading task with the Times New Roman font only. The individual reading curves are shown in Figure 2-4. Reading speed was modelled as a function of both print size and character count (Eq.3, see method), and individual reading indices are listed in Table 2-1.

The MRS varied widely across the low vision participants, ranging from 30 to 398 words per minute. CPS also had a wide distribution, ranging from 0.47 to 1.84 logMAR. These wide individual differences are not surprising, given the heterogeneity of the low-vision sample. Despite these substantial differences in overall reading ability, the CCC was similar across display formats, averaging 8.4 characters. There was a slightly smaller value in the tablet format, which was 1.8 characters smaller than the laptop ($p < 0.001$, 95% CI = [1.0,2.6]), and 1.5 characters smaller than the phone ($p < 0.001$, 95% CI = [0.6,2.4]). Again, there was no significant correlation between CCC and CPS for any of the displays (all $p > 0.05$).

For some low vision participants and display formats, the CPS was larger than the print size corresponding to the CCC, meaning that there was no print size to achieve the criterion of 80% of MRS. For example, subject LV10 had similar CPS values of approximately 1.4 logMAR across the three display formats, but for the phone format there was no print size above this critical size that allowed 8 characters to be fit on each line (Figure 2-4).

The constrained-MRS values were smaller than the MRS for all three displays, by 7.6% for laptop ($p = 0.022$, MD = 12.9 wpm, 95% CI = [2.4, 23.4]) and 6.2% for tablet ($p = 0.032$, MD= 10.6 wpm, 95% CI = [1.2, 20.0]). The phone display showed the largest reduction of 20.7% ($p = 0.005$, MD = 37.2 wpm, 95% CI = [14.3, 60.0]).

Discussion

We have defined the concept of “critical character count” (CCC) representing the minimum number of characters per line to achieve a criterion of 80% of an individual’s maximum reading speed. We have shown how the critical print size required for this near-maximum reading speed interacts with the critical character count to constrain reading performance. Our findings are relevant to the usefulness of small displays for people with normal vision and the requirements of display format for people with reduced acuity.

Three major findings emerge from the results: 1) The CCC is constant across fonts and display formats. On average, it is 13 characters for normally sighted participants and 8 characters for low vision participants. 2) The range of print sizes to achieve the near-maximum reading speed has a lower bound determined by the CPS and an upper bound determined by the CCC. When the CPS is greater than the print size associated with the CCC, no print size will support the near-maximum reading speed. 3) Even within the range of print sizes limited by the lower and upper bounds, the highest achievable reading speed (constrained-MRS) will often be less than the unconstrained maximum reading speed (MRS).

Earlier studies have investigated the impact of window size on reading, in the context of magnifiers for low vision (summarized by Legge, Chapter 4, 2007). Window size is the number of characters visible on a line of text in a magnifier's field of view, which decreases when the power of the magnifier increases. There are two major differences between our current investigation on critical character count and the earlier studies on critical window size. First, in the current study, the constraints on visible text were imposed by the display format rather than by the field of view of a magnifier. Second, we allowed the print size and character count to covary as they would for any fixed-size display, rather than controlling one factor and varying the other. This approach allowed us to investigate the joint effects of print size and display format on reading in a more realistic context.

The current study has identified a critical value for the number of characters per line (CCC) as a fundamental limitation on reading speed. What accounts for the critical value? The CCC value of 13 for normal vision is consistent with the possibility, as outlined in the Introduction, of a line-length limitation due to perceptual span. Estimates of the size of the perceptual span to the right of fixation in normal vision of 10 to 15 characters are consistent with this possibility (Rayner, 1986). Lending additional support, previous findings showing that the perceptual span is smaller in low vision (16-18) are consistent with our finding that the CCC is lower in low vision. It has also been shown that perceptual span appears to be constant across fonts (Rayner et al., 2010).

The difference in CCC between normal and low vision might also be related to the time required for the eyes to retrace from the end of one line to the beginning of the next line (Rayner, 1998; Parker et al., 2019). When the number of characters per line decreases, the number of these return sweeps (and number of lines to be read) increases, taking up proportionately more time in reading. The impact on reading speed would be greater for fast readers with normal vision (as the associated time cost would be proportionally larger) than slower readers with low vision. This difference might contribute to a larger CCC in normal vision.

We also considered the possibility that the CCC is related to the distribution of word lengths in text. About 80% of English words in text are 9 or fewer letters (Computed from the frequency distribution of word lengths in the Corpus of Contemporary American English, 2020), thus a CCC of 8 or 9 letters could avoid most word splitting across lines. This observation seems consistent with the CCC of 8 we found for low vision, but not the higher value of 13 for normal vision. Taken together, our findings are most compatible with the perceptual span as the primary determiner of the CCC.

What are the real-life implications of our findings? The CPS for people with normal vision is approximately 0.2 degrees (equivalent to 0.38 logMAR) (Legge & Bigelow, 2011). If text is presented on a smart watch at a viewing distance of 40 cm for a normally sighted user, the minimum display width would need to be about 1.7 cm in order to include 13 characters per line at the required print size.

For reading with low vision, our findings have implications for two groups: eye-care clinicians and display designers. Clinicians may wish to recommend digital displays for their patients. How large should the display be? An individual's CPS can be measured directly from a test such as MNREAD (Legge et al., 1989) or estimated from a measure of letter acuity (Xiong et al., 2018). Our findings establish the relationship between CPS and the minimum display size required for the critical number of characters per line. Figure 2-5 plots minimum display width as a function of letter acuity (lower axis) or CPS (upper axis). The blue line represents a standard reading distance of 40 cm and the red line a reading distance of 15 cm. Low-vision readers often adopt shorter than normal viewing distances for reading but rarely less than 15 cm (Granquist et al., 2018) because of difficulties of posture or accommodation. The equation generating the red and blue lines is provided in Appendix 2. These lines show the results for Times New Roman, while the surrounding gray bands represent the corresponding values for eighteen fonts (see Appendix 4 for details). The variation across most fonts is tiny. This is because of a fortuitous tradeoff: fonts with wider character spacing such as Courier take up more horizontal space but have a smaller CPS (Xiong et al., 2018). This means that Figure 2-5 provides guidance for display selection and is appropriate for a wide range of fonts that may be encountered by the low-vision reader. The intersection of the vertical dashed lines and the blue and red lines show the minimum display widths for low-vision readers with three levels of acuity. For a low vision reader with visual acuity of 1.0 logMAR (the boundary for legal blindness in the United States), the minimum display width for near-maximum reading is 11.0 cm at 40cm viewing distance (intersection with the blue line), which excludes use of a smart phone with a width of 10 cm (Landscape). However, if the person is comfortable reading at 15 cm (intersection with the red line), this smart phone now meets the minimum display width of 4.2 cm.

Display designers can use Figure 2-5 to estimate the inclusiveness of their devices for users with low vision. For a display of a given width, the intersection of a horizontal line at this display width with the blue and red lines in Figure 2-5 show the

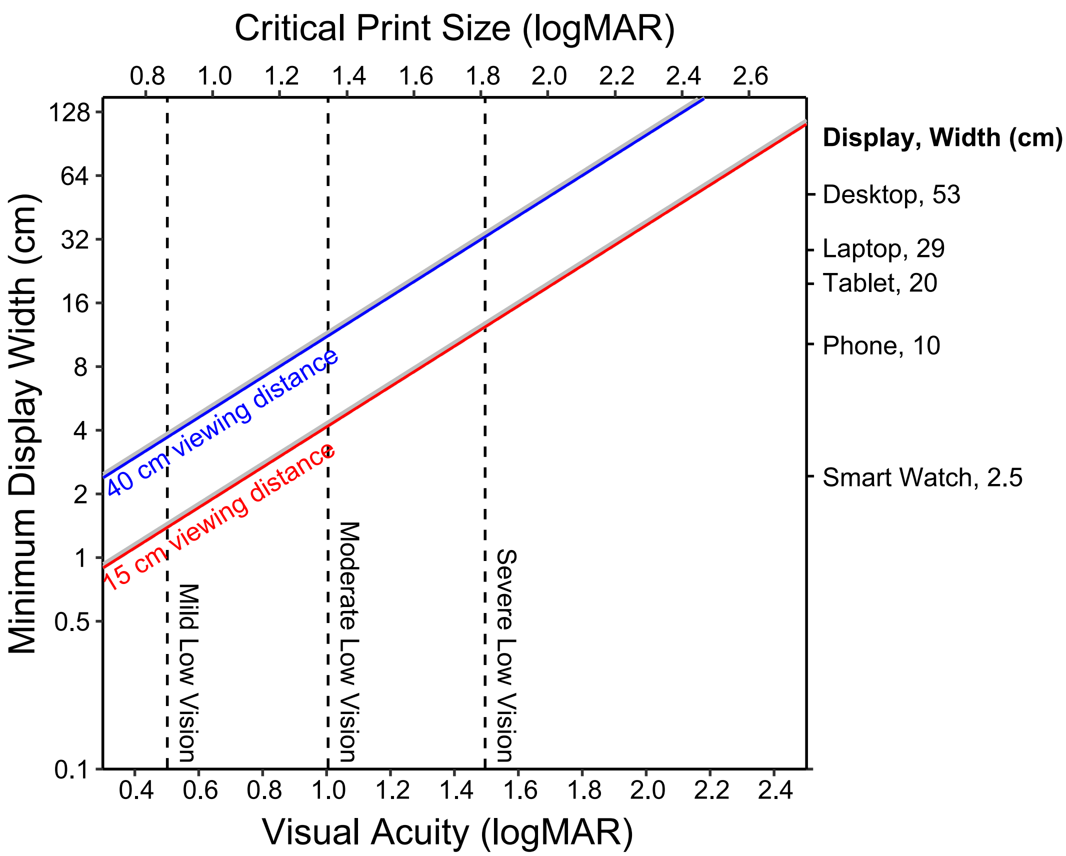


Figure 2- 5: Minimum display width for low vision individuals. The minimum display widths are presented as a function of visual acuity (VA, bottom x-axis) for two viewing distances (40 cm, blue line: y = 1.22*100.96VA and 15 cm, red line: y = 0.46*100.96VA), where y is the minimum display width in cm and VA is visual acuity (logMAR). The derivation of the equations for the red and blue lines are provided in Appendix 2. These equations use a liberal CCC value of 9 characters, rounding up to the nearest integer value from the average value of 8.4 found in our experiment for low vision. The critical print sizes (CPS) corresponding to each VA are shown on the top x-axis. The relationship between VA and CPS was obtained from 87 low vision participants (Xiong et al., 2018; Cheong et al., 2008; Calabrese et al., 2018), a scatter plot is provided in Appendix 3). Eighteen fonts* were included in the analyses (Appendix 4), with the colored lines showing results for Times new Roman and the gray ribbons representing the results across all fonts. Dashed vertical lines show examples for Mild, Moderate and Severe low vision with corresponding acuities of 0.5, 1.0 and 1.5 logMAR. The intersections of these vertical lines with the blue and red diagonal lines represent the minimum display widths for these acuities and reading distances. The display widths of five common digital displays (smart watch, phone, tablet, laptop, and desktop computer) are shown on the right y-axis.

* The eighteen fonts include five common fonts (Times, Courier, Helvetica, Arial, Calibri) and their italic and bold variations and three new fonts developed for patients with macular degeneration (Eido, Eido Mono (Bernard et al., 2016) and Maxular (Letters between friends, 2017)). Font details are provided in Appendix 4.

required acuities to achieve the critical number of characters per line for viewing at 40 cm (blue line) or 15 cm (red line.) The larger the logMAR acuities required, the more inclusive is the device for low vision.

Table 2-3 shows the acuity requirements for the five display sizes listed on the right of Figure 2-5. For example, a smart phone with landscape width of 10 cm can accommodate users with visual acuities up to 0.97 logMAR at 40 cm viewing distance and 1.41 logMAR at 15 cm viewing distance. (The corresponding Snellen fractions are close to 20/200 and 20/500.

Display	Width (cm)	Required Acuity at 15 cm distance Mean [Min, Max] (logMAR)	Required Acuity at 40 cm distance Mean [Min, Max] (logMAR)
Desktop	53	2.16 [2.14, 2.17]	1.71 [1.70, 1.72]
Laptop	29	1.88 [1.87, 1.89]	1.43 [1.42, 1.44]
Tablet	20	1.71 [1.70, 1.72]	1.27 [1.25, 1.28]
Phone	10	1.41 [1.40, 1.42]	0.97 [0.95, 0.98]
Smart Watch	2.5	1.27 [1.25, 1.28]	0.76 [0.75, 0.77]

Table 2- 3. Acuity requirements for five digital displays viewed from 40 cm or 15 cm. The acuity requirement is determined by the angular size of letters when there is an average of nine characters per display width.

In this article, we have provided a unified framework for understanding the interacting effects of print size and display format on reading speed. Not only must characters exceed a critical size for near-maximum reading, there must also be more than a critical number of characters per line (CCC). Our analysis reveals a requirement of about 13 characters for normally sighted readers, and about eight characters for people with low vision. Our findings have implications for the design of small text displays for people with normal vision and the prescription of appropriate reading aids for people with low vision.

Chapter 3: Crowding and Parallel Visual Pathways

Introduction

A central function of the visual system is identifying objects in our environment. Often the objects are in the visual periphery and surrounded by other items. While recognizing and discriminating objects in isolation is an easy daily task, it becomes substantially harder when objects are cluttered. Crowding is described as the detrimental effects of nearby items on visual recognition of objects (Whitney, & Levi, 2011). Especially in peripheral vision, people's ability in recognizing even simple objects is remarkably impaired by the neighboring objects regardless of the category of those objects. Given that objects are often in our visual periphery and rarely in isolation, crowding is considered as a bottleneck of object recognition (Levi, 2008). Although crowding can be observed with numerous different visual stimuli including letters, faces, moving vs. static objects, with low or high spatial frequencies, in low or high contrasts, a prominent example of crowding's influence in daily life is in the domain of reading. Many studies investigated the impact of crowding on reading (Pelli et al., 2007; Levi, Song, & Pelli, 2007; Chung, 2007; Martelli, Di Filippo, Spinelli, & Zoccolotti, 2009). Taken together, understanding the mechanisms of crowding effect is crucial to have a comprehensive understanding of processes in reading and object recognition.

The crowding phenomenon has been studied since the 1930s (Ehlers, 1936). A landmark research from Bouma (1970) showed that for visual recognition of a target at Øº eccentricity, the nearby objects (flankers), need to be at least Ø/2º away from the target in order to prevent the crowding effect. A number of studies also revealed key properties of crowding based on the location of the target and flankers in the peripheral visual field. Flankers placed along the fovea-target radial direction impair the recognition performance more than on the tangential direction (radial/tangential anisotropy; Toet & Levi, 1992), and the far flanker (flanker more eccentric than the target) is more detrimental than the near flanker (flanker closer to the fovea than the target) (inward-outward anisotropy; Bouma, 1970). It was also shown that crowding is stronger in the upper than lower visual field (He, Cavanagh, & Intriligator, 1996) despite the fact that the area of cortical representation of upper and lower visual fields is roughly the same in the primary visual cortex. Furthermore, the effect of number of flankers (Pelli, Palomares, & Majaj, 2004) and whether they could be grouped (Manassi, Sayim, &Herzog, 2012) affect crowding. In addition, stronger crowding is observed with higher target-flanker similarity (Kooi, Toet, Tripathy, & Levi, 1994).

In contrast to the extensive characterization of the spatial properties of crowding, there is a lack of understanding of the neural mechanisms of crowding. Previous studies have often focused on investigating the stage(s) in the visual hierarchy where crowding starts to limit target processing (Pelli, 2008; Anderson, Dakin, Schwarzkopf, Rees & Greenwood, 2012; Millin, Arman, Chung & Tjan, 2014; Chicherov, Plomp & Herzog, 2014). While there is no consensus on a specific locus of crowding in the visual system, it has been proposed that crowding may occur at multiple stages of visual processing (Louie, Bressler & Whitney, 2007; Fischer & Whitney 2011; Anderson et al., 2012; Manassi & Whitney, 2018). Fischer and Whitney (2011) argued that crowding does not stop object processing in a particular point in the visual system or that objects is not broken down to their low-level features. Theories for neural mechanisms of crowding range from proposals emphasizing the role of receptive fields and hypercolumns to proposals addressing more high-level processes including feature integration and attentional processes (Levi, Klein, & Aitsebaomo, 1985; He et al., 1996; Pelli et al., 2004; for a review, see Levi, 2008). Importantly, there is little consideration on the role of parallel visual pathways and the potential difference of crowding among them.

From the retina to the primary visual cortex, visual information is processed along two major channels, the Magnocellular (M) and the Parvocellular (P) pathways. Information processed in the primary visual cortex is further processed along the ventral "perception" and the dorsal "action" pathways, with the dorsal stream receiving more input from the M pathway and the ventral stream receiving more input from the P pathway (Nassi & Callaway, 2009). Neurons in these two channels differ from each other in terms of both their anatomy and corresponding functions. Many studies have shown that the two visual pathways are tuned to a number of distinct visual features. For example, the P pathway is very sensitive to color whereas the M pathway is blind to color (Livingstone, & Hubel, 1987). The P pathway favors spatial details while the M pathway cannot resolve high spatial frequencies (Derrington and Lennie, 1984). On the other hand, the M pathway is very fast at processing visual information, while the P pathway is relatively slow. In relation to these distinctive features of the two pathways, different visual phenomena have been investigated to uncover whether one pathway is more involved compared to the other in daily visual functions. To name a few, researchers have looked at binocular rivalry (Carlson & He, 2000), dyslexia (Stein, 2001) and consciousness (Tapia & Breitmeyer, 2011). However, no published research has reported the relationship between crowding and parallel visual pathways.

There are a large number of studies investigating the effects of stimulus properties in crowding (Tydgat, & Grainger, 2009; Grainger, Tydgat, & Issele, 2010; Pelli, 2008). Surprisingly, given that visual parallel pathways represent biases towards different stimulus properties and thus these neural pathways are differentially involved in processing different types of stimuli, little consideration was explicitly given to the possibility that the nature of the crowding effect could be different across these visual pathways. For example, a recent study showed that crowding effect is dissociable in color and motion processing, which allowed the authors to make the point that crowding is not a singular process (Greenwood & Parsons, 2020). However, color and motion are stimulus properties that are biased to be processed in the Parvocellular and Magnocellular pathways. To further our understanding, we ask the question how these biases in the feature dimensions influence crowding. Uncovering these possible pathway differences will help us to better understand the neural correlates of crowding.

To our knowledge, this study is the first in the field that explicitly investigated the potential differences of the crowding effect in different visual pathways. We studied the vulnerability to crowding in the two parallel pathways with stimuli designed to selectively engage the P pathway or the M pathway. In Experiment 1, we aimed to isolate the two pathways by using biased stimulus for the targeted pathway and backgrounds to saturate the other pathway. In the Experiment 2, we used different stimuli and tasks to investigate crowding properties for form, color, and motion discrimination. In both experiments, we measured the critical spacing of crowding, the minimal distance between target and flanking objects that allows identification of the target object, a smaller critical spacing would indicate weaker crowding effect.

EXPERIMENT I

Experiment 1 investigated the spatial properties of crowding with stimuli designed to separately engage the Parvocellular or Magnocellular pathway, by tuning stimulus features to activate the targeted pathway and manipulating background to saturate the other pathway. The same target discrimination task was used for all stimuli.

Methods

Participants

Thirteen undergraduate students from the University of Minnesota, aged between 18-30 years, with normal or corrected vision participated in the study. All participants were recruited from the University of Minnesota Psychology Department’s Research Experience Program participant pool, and gave written informed consent to participate in accordance with the policies approved by the human subjects review committee of the University of Minnesota.

Apparatus and Stimuli

The stimuli were generated using MATLAB with the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997) and displayed on a 24-inch TOBI T60XL screen (refresh rate: 60Hz, resolution:1920 x 12000). The monitor was calibrated using a Photo

Research PR-655 spectrophotometer, luminance gamma curves were measured and inverted with a look-up table.

Achromatic contrast discrimination and pulsing stimuli have been used to assess the M and P pathways' functions in psychophysical experiments (Pokorny & Smith, 1997; Leonova, Pokorny & Smith, 2003; McAnany & Levine, 2007). A similar paradigm was developed in the current experiment to test crowding effect in parallel pathways. In the P pathway biased condition, the background was filled with fast flickering (30 Hz) small red squares, constantly changing in their luminance. The target and flankers were defined as L-shaped 3-squares, in green color and fast flickering, so that the luminance modulation was similar to the background (Figure 3-1A). The mean luminance of each square was equal to the mean luminance of the background. In the M pathway biased condition, the background was filled with isoluminant green and red squares, and each square was 0.5º x 0.5º in size. The target and flankers were defined as three squares in the periphery, forming an "L" shape. They differed in their luminance modulation (Figure 3-1B) from the background and these L-shaped 3-squares flickered at a fast rate (30Hz) to be distinguished from the stable background. In both conditions, in each trial, the target and flankers were randomly assigned to one of the four possible orientations (Figure 3-1C), with the possibility of the target and flankers having the same orientation. Two flankers were aligned horizontally to the target, with one flanker on each side (of the target). The stimuli were presented for 200ms and the eccentricity of the target was 9º in the radial direction with respect to the fixation point.

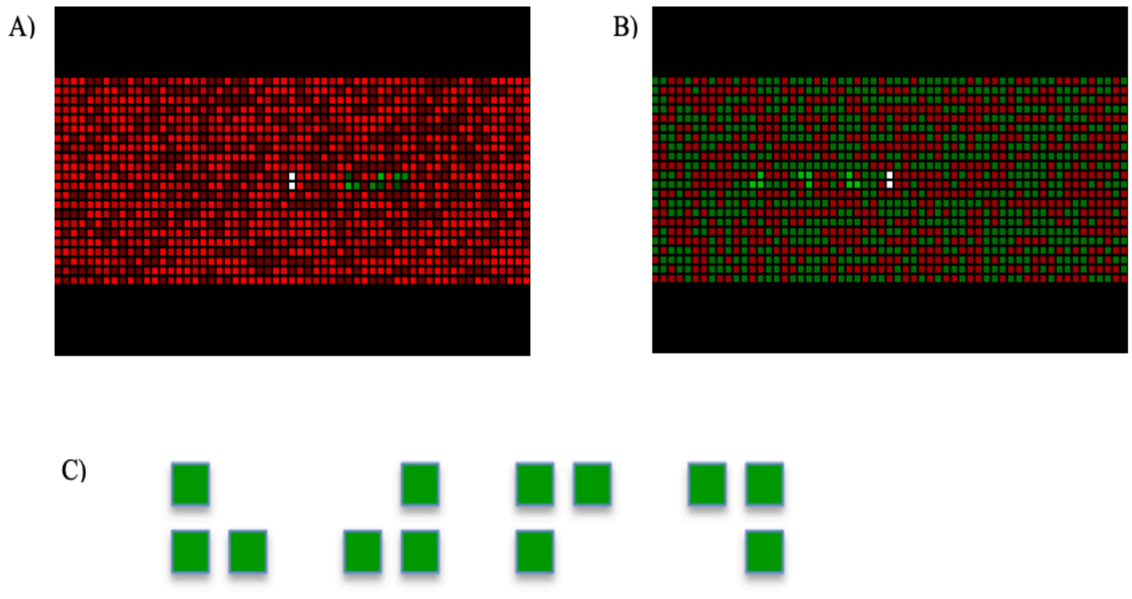


Figure 3- 1:Stimuli for Experiment 1. A) P pathway biased condition. Stimuli defined by green color and background defined by red color. Both the stimuli and background constantly flickering at a fast rate, by changing in luminance. B) M pathway biased condition. Stimuli defined by luminance flicker while the background is isoluminant, stable, green-red colored. C) Possible orientations for the target and flankers in Exp1.

Procedure

The viewing distance was 60 cm and participants' heads were stabilized with a chinrest. The middle two squares in the background were presented in white color to serve as a fixation point. Participants were instructed to look at these two white squares throughout the whole experiment.

Before the main crowding experiment, participants were asked to complete a minimal flicker procedure. This procedure was designed to find the subjective isoluminance values for red-green stimuli for each participant. Participants were asked to fixate their eyes on the black dot in the middle of the screen and adjust the luminance of a square located in their peripheral vision that was presented in the right or left of the black dot at 9º eccentricity. Participants used the arrow keys on the keyboard until the flickering was minimal. This procedure was repeated 8 times, 4 times on each side of the screen. Four different green values were obtained from this procedure for each participant as isoluminants of given red values. The luminance of red squares were one of the four values; 11.46, 15.94, 20.43 and 24.91 cd/m^2. Only these sets of four red and green values were used to render the background and the stimuli in the experiment.

Eye-tracking was performed both during the minimal flicker procedure and main experiment to make sure participants maintained fixation. The screen-based eye-tracker TOBII T60XL was used and data were recorded binocularly. Participants were calibrated using a standard 9-point grid.

In the main experiment, participants were instructed to perform a peripheral orientation discrimination task. While participants fixate at the fixation point in the middle of the screen, target and flankers briefly appeared either right or left side of the fixation. After the presentation of the stimuli, the word “response” was presented, prompting the participant to report the orientation of the target “L” shape. Participants responded by pressing the corresponding key for the particular orientation. They were instructed to respond in two seconds following the target/stimuli presentation, and the next trial was presented immediately following the response. All participants completed two conditions, each aiming to target either M or P visual pathways. The order of the conditions was counterbalanced. Each condition consisted of 6 blocks, varying in their inter-stimulus spacing, in other words, the distance between the flankers and the target. The spacing in the blocks were 2.5º, 3.3º, 4.2º, 5º, 6º, and one baseline block (0º) presenting no flankers. Each block consisted of 96 trials, as a total of 576 trials for each pathway condition. Critical spacing for each condition was defined as the distance where participants were able to reach 80% identification accuracy.

The accuracy performance was fitted as a function of target-flanker spacing with a cumulative Gaussian sigmoid curve using the Psignifit toolbox software for MATLAB (Wichman and Hill, 2001).

Results and Discussion

Consistent with prior research (Bouma, 1970; Greenwood, Bex & Dakin, 2010; Toet & Levi, 1992; Pelli et al., 2004) target identification performance improved as target-flanker spacing increased in both M pathway and P pathway conditions. Figure 3-2A shows the proportion of correct responses to different target-flanker spacings for both conditions. Average performance in the P pathway biased condition was better throughout the whole inter-stimulus spacing range. The critical spacing for each condition was defined at 80% correct identification.

Figure 3-2B shows the critical spacing for each individual participant along with the averages for each condition. We found that the critical spacing in the M-biased condition (M = 5.76º) was significantly larger compared to that of the P-biased condition (M = 4.49º) (paired-t (12) = 4.3, $p<0.001$), and that this difference was consistent across all subjects, despite the individual differences in critical spacing values. In both experiments, participants' performance was around 90% in non-crowded (baseline) condition, and around 40-45% in the most difficult crowding condition (2.5º inter-stimulus spacing), suggesting that the difference in critical spacing between the two conditions was not due to possible differences in task difficulty.

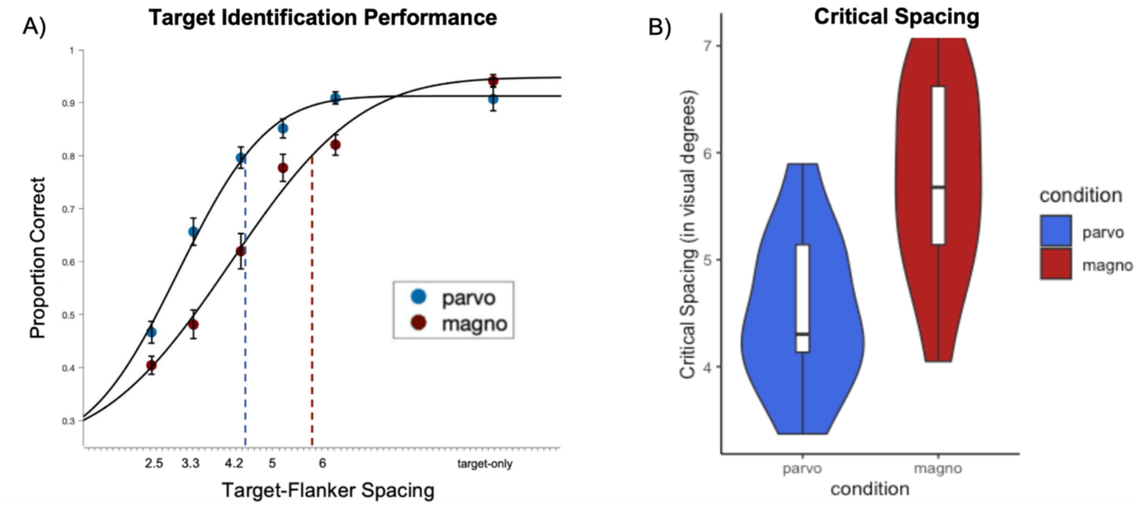


Figure 3- 2: Results of Experiment 1. A) Target identification accuracy as a factor of target-flanker spacing. Two curves for M pathway biased (magno) and P pathway biased (parvo) conditions, fitted with psychometric function. The dashed lines indicate the value of critical spacing for the two conditions, corresponding to 80% accuracy. B) Violin plot of critical spacing data for both Magnocellular pathway (magno) and Parvocellular pathway (parvo) conditions. The width of the bar represents the density of distribution of the individual data. The white bar indicates the interquartile range with the median line, and 95% interval is shown with the black vertical line.

Results of Experiment 1 indicate that the P pathway compared to M pathway can better resist crowding in form perception, or that the crowding effect in this experiment is more severe in the M pathway than in the P pathway. Our results do not suggest specific differences in the underlying crowding mechanisms in different visual pathways, and instead may reflect that at the same eccentricity, neurons in the P pathway tend to have smaller receptive fields than that in the M pathway (Dacey & Petersen, 1992; Nassi & Callaway, 2009).

The goal of the first experiment was to isolate the two pathways and study their spatial properties in crowding, however, in real life, visual objects are rarely, almost never, completely fit the description of P pathway or M pathway stimulus. Therefore, in Experiment 2, we investigated crowding using stimuli that are in different functional categories and are more relatable to real life, yet with clear link to the idea of two parallel pathways.

EXPERIMENT II

This experiment examines the question of whether recognition performance in crowding yields different patterns depending on the parallel visual pathways, by applying different stimuli categories and their corresponding visual tasks. Three different stimuli and their corresponding tasks were used to investigate function-specific crowding properties. Specifically, color (primarily P pathway), motion (more M pathway than P pathway) and form (more P pathway than M pathway) discrimination tasks are implemented.

Methods

Participants

Twenty-five undergraduate students from the University of Minnesota, aged between 18-30 years, with normal or corrected vision participated in the study. None of the participants from the first study were included in this experiment. All participants were recruited from the University of Minnesota Psychology Department's Research Experience Program participant pool. They gave written informed consent to participate in accordance with the policies approved by the human subjects review committee of the University of Minnesota.

Apparatus and Stimuli

As in Experiment 1, the stimuli were generated using MATLAB with the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997) and displayed on a 24-inch TOBI T60XL screen (refresh rate: 60Hz, resolution:1920 x 1200).

In the motion condition (targeting the M pathway), the target and flankers were 2.5º squares, filled with isoluminant red/green vertical bars with the spatial frequency of 1.6 cyc/deg and luminance of 18.9 cd/m^2. (Figure 3-3A). The chromatic gratings had the spatial frequency of 0.5 cyc/deg, was modulated sinusoidally, and moved to the right or left direction, alternating randomly. The final stimulus was essentially an achromatic sine wave grating moving from the right or left, overlaid above the chromatic gratings, simulating movement across the red/green gratings (Figure 3-3D). This motion-on-color paradigm was adapted from Wen et al. (2015). The temporal frequency of the luminance grating was 15 Hz. Based on previous evidence that the P pathway is consumed by the high-contrast color modulation and that the detection of motion of the low-contract luminance grating is controlled by the M pathway, this task aimed to target the M pathway (Merigan, Byrne & Maunsell, 1991).

The novel experimental paradigm in the color condition was developed based on previous studies that tested chromatic information in the periphery in relation to parvocellular pathway (Lee, Pokorny, Smith, Martin & Valbergt, 1990; Cooper, Sun & Lee, 2012). In this condition (targeting the P pathway), the background was gray with a black fixation dot (1º) in the middle of the screen. The target and flankers were defined as squares in the periphery, 2.6º in size. Target squares were presented as a of a single shade of yellow, whereas the flankers were divided into four quadrants, each with a different shade of yellow (Figure 3-3B). The luminance of target and flankers were 18.9 cd/m^2 and their hue values were individually determined for each participant. Given the fact that the P pathway is highly responsive to color information while the M pathway is insensitive to it (Schiller, Logothetis & Charles, 1990), this task almost exclusively engaged in P pathway.

Lastly, in the form condition, the target and flankers consisted of a green vertical bar and a red horizontal bar, forming a 2.5º cross shape (Figure 3-3C). In all trials, the red horizontal bar was overlaid on top of the green vertical bar at the intersection, and the two bars were isoluminant (luminance: 18.9 cd/$m^{2)}$. This line displacement task was chosen among possible form discrimination tasks in order to equate the task difficulty with the motion and color discrimination tasks. This task allows 2AFC response structure with task difficulty flexibly adjusted for the participants. The aim of this task was to target both parallel visual pathways. Based on the evidence that form discrimination is primarily performed in the P pathway along with some information carried by the M pathway (Livingstone & Hubel, 1987), this task was expected to engage both visual pathways with an emphasis on the P pathway.

In all three conditions, two flankers were always aligned horizontally to the central target. The stimuli were presented for 250 milliseconds and the eccentricity of the target was 11º in the radial direction with respect to the fixation point.

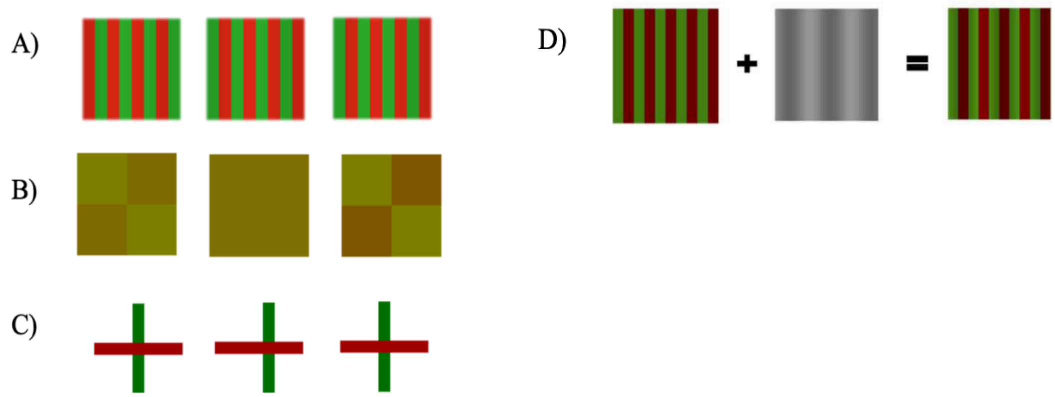


Figure 3- 3: Three stimuli sets for Experiment 2. A) Motion crowding condition. Motion that is obtained by the luminance change of the bars is not shown in the figure. B) Color crowding condition. C) Form crowding condition. In all conditions, middle object is the target with two flankers in each side. D) The motion-on-color paradigm, figure adapted from Wen et al., 2015.

Procedure

The procedure was very similar to Experiment 1. In this experiment the fixation point was a black dot (visual angle of 1º).

Before the main crowding experiment, as in Experiment 1, participants were asked to complete the minimal flicker procedure. In addition, participants also completed a color threshold test. This test allowed researchers to obtain subjective color thresholds for each participant. In this procedure, similar to the minimal flicker procedure, participants fixated on the white dot in the middle of the screen and adjusted the hue of a square that was presented in the periphery at 11º. Participants were instructed to adjust the hue until they find the “perfect” yellow. As they changed the hue of the square using the arrow keys on the keyboard, the luminance of the square stayed constant. This trial was repeated 6 times, 3 times on each side of the screen. The starting hue value in each trial was randomized to avoid biases. The average of the 6 trials was calculated to determine the subjective “yellow” value and was implemented in the color condition of the main experiment.

Given the success of the participants in the first experiment in keeping their gaze at the fixation point, eye-tracking was not used in this experiment.

In the main experiment, each condition had a specific task based on the stimuli used. In the color condition, participants were asked to decide if the hue of the target square was closer to green or red color, compared to their internal “perfect” yellow representation. As noted above, the target stimulus was manipulated based on each participant’s subjective result in the color threshold procedure. Target stimuli were created by adding/subtracting 3, 5, 7, 9 steps to/from the “perfect” value (i.e. if a participant’s average value from color threshold test is [128 120 0], the manipulated colors for step 3 would be [128 123 0] or [128 117 0]). The order of the different steps of hue manipulation was counterbalanced in each block. Flankers were divided into quadrants and each quadrant had a different hue value. These values were, again, based on each participant’s subjective results, by subtracting/adding 0, 10, 20, 30 steps. The representation of different hues in different corners of the main square was randomized. The response procedure was the same as Experiment 1 except that participants were instructed to press the “up” arrow key, if they think the target stimulus’ hue is closer to green, and the “down” arrow key if it is closer to red.

In the motion condition, the task was deciding the direction of the “moving” wave, which was created by applying a luminance sine-wave grating, perceived over the chromatic gratings. The direction was defined either to the left or to the right both in the target and flankers. The direction of the moving wave was chosen randomly, with the possibility that one or both of the flankers and the target could move in the same direction. Participants were instructed to press the “right” or “left” arrow key in the response period immediately following the presentation of the gratings, based on their perception of the moving sinusoidal grating.

In the form condition, participants were asked to decide whether the red horizontal bar was longer in the right or left side of the green vertical bar in the global cross sign shape. The horizontal bar was manipulated to be shifted to right or left 0.25º for target and 0.25º or 0.35º for flankers. The side (right or left) and the degree (for flankers) to which the horizontal bar was shifted were randomized in each trial for both the target and the flankers, allowing for the possibility of the target and one or more flankers to have the same direction and/or degree of shifting. Participants responded by pressing the “right” or “left” arrow key, indicating the longer side of the horizontal bar in the target.

All participants completed all three conditions, and the order of the conditions was counterbalanced. Each condition consisted of 4 blocks, varying in their inter-stimulus spacing. As in Experiment 1, one of the blocks was defined as the baseline, as there was no flanker presented. In addition to the baseline block (0º), the spacing in the other 3 blocks were 3.4º, 5.6º and 7.7º. Each block consisted of 56 trials, with a total of 224 trials for each condition.

Seven participants were later excluded from the analyses as they did not reach 80% accuracy in the target-only (noncrowded) block in at least one of the three conditions. High level performance in the target-only condition indicates that participants can perform object recognition tasks in their periphery in the absence of distractors.

Results and Discussion

Like in Experiment 1, recognition performance improved as the target-flanker spacing increased. Figure 3-4A demonstrates the target identification performance as a function of target-flanker spacing in the normalized data. Data was normalized to fix the target identification performance at 90% accuracy at the target-only block for all three conditions.

The rate of improved target identification with increasing target-flanker spacing differed across conditions. For example, form discrimination, which is the most commonly used task in crowding studies (presumably engaging both P and M pathways), showed the most severe crowding effect and considerable benefits of large target-flanker spacing. On the other hand, in the color discrimination task (presumably biased towards the P pathway), minimal crowding effect and little benefit from larger target-flanker spacing were observed. We quantified the crowding effects using critical spacing, as described in Experiment 1. The average critical spacing value was 3.72º for the color condition, 5.98º for the motion condition and 6.56º for the form condition. Figure 4B shows the critical spacing for each individual participant along with the averages for each condition. A repeated measures ANOVA showed a significant difference in critical spacing among conditions ($p<0.05$). Post-hoc pairwise comparisons revealed that the difference between color-form conditions was significant ($p<0.01$).

These results can be interpreted such that form discrimination, which requires spatial integration of multiple features of an object, is more vulnerable to crowding than motion and color discrimination, which are based on a single property of the target object and do not involve spatial integration. Thus, current results support the idea that feature integration plays an important role in crowding effect (Pelli, Palomares & Majaj, 2004). Additionally, our results showed that color discrimination of a flanked target is much less affected by the typical crowding effect.

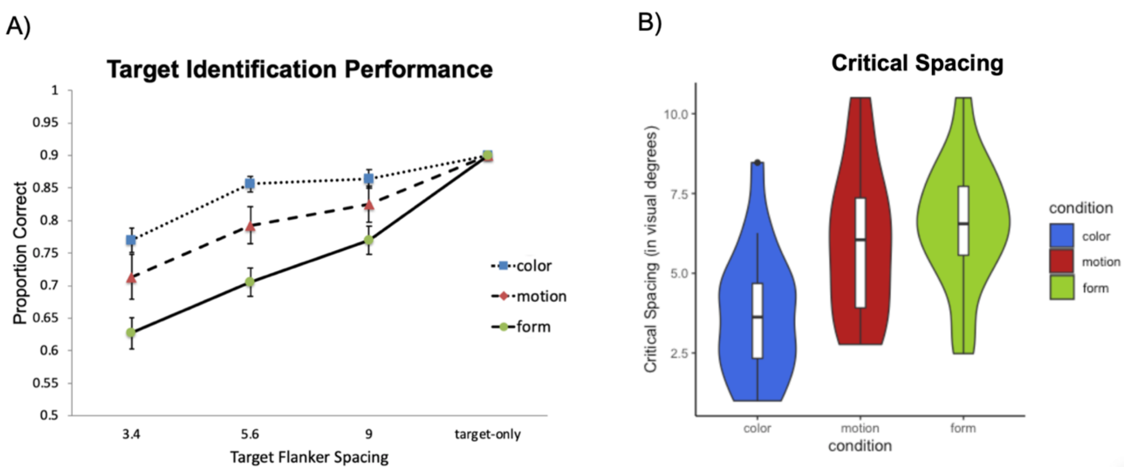


Figure 3- 4: Results for Experiment 2. A) Target identification performance as a factor of target-flanker spacing for all three conditions. Data is normalized to fix the target-alone condition at 0.9. B) Violin plot of critical spacing data for color, motion and form conditions, respectively. The width of the bar represents the density of the distribution of the individual data. The white bar indicates the interquartile range with the median line, and 95% interval is shown with the black vertical line.

In everyday life, the visual system detects and discriminates objects in the periphery that are moving, have certain colors, and possess forms composed of multiple features. The results of the current experiment further our understanding of how motion, color, and form differentially relate to crowding, and provides quantitative measures of crowding effects for these functional important and ecological valid properties.

General Discussion

Understanding the neural correlates of crowding is important, given that crowding is considered to be a primary limiting factor for object recognition in the periphery (Levi, 2008). Crowding is proposed to occur at multiple stages of visual system (Louie et al. 2007; Anderson et al., 2012; Manassi & Whitney, 2018), with neural correlates identified from retinal ganglion cells (RGCs) to higher-level object recognition mechanisms (Levi, Klein, & Aitsebaomo, 1985; Herzog & Manassi, 2015; Kwon & Liu, 2019). The current study investigated whether spatial properties of crowding vary in two major visual pathways where the information is processed in parallel in the visual system. Experiment 1 demonstrated that, at a particular eccentricity, the P pathway is spatially more resistant to the crowding effect compared to the M pathway. One explanation for this difference is that the parasol ganglion cells, which project to the magnocellular layers of the lateral geniculate nucleus (LGN), have larger receptive fields than the midget ganglion cells which project to the parvocellular layers of LGN (Rodieck, Binmoeller, & Dineen, 1985; Dacey & Petersen, 1992; Dacey, 2000). For stimuli that are processed mainly in the M pathway, the spatial extent of the information integration is larger, which may lead to stronger crowding effect compared to the stimuli mainly processed in the P pathway. Additionally, in a recent study, Kwon and Liu (2019) showed that the spatial asymmetries in crowding (e.g. radial/tangential anisotropy, inner vs. outer flanker effect) could be explained by the sampling density of RGCs across the visual field. Although Kwon and Liu (2019) did not differentiate between parasol and midget cells, their study, together with known properties of these cells (Dacey, 1994), provide support for our hypothesis that the characteristics of crowding effect may vary across the parallel visual pathways starting from an early stage.

We also investigated the relationship between the spatial properties of crowding effect and the two visual pathways in the context of higher-level visual processing by using three different tasks. Our findings illustrate that the crowding effect is not uniform across different visual tasks that differentially engage the two major visual pathways. In the form discrimination task, the crowding effect was more sensitive to target-flanker spacing; whereas color discrimination task was relatively insensitive to target-flanker spacing with an overall weaker crowding effect. Interestingly, the motion discrimination task showed intermediate level of sensitivity to target-flanker spacing. We note that in both experiments, color defined stimuli were less affected by crowding compared to targets defined by luminance flicker or motion. Form discrimination apparently was most vulnerable to crowding, possibly because it requires integration of multiple features and involves both the M and P pathways.

Our findings provide additional evidence for non-uniformity of the crowding effect as suggested by a number of previous studies (Bex & Dakin, 2005; Kennedy & Whitaker, 2010; Greenwood & Parsons, 2020). Although most previous studies focused on form crowding (Pelli, Palomares & Majaj, 2004; Lev, Yehezkel & Polat, 2014; Tripathy, Cavanagh & Bedell, 2014; Agaoglu & Chung, 2016), it has been demonstrated that we need to consider the differences in the nature of stimuli and tasks to fully understand crowding. In a study of motion crowding using non-static Gabor patches, it was shown that theories such as compulsory averaging, developed to account mostly for form discrimination, must be modified to capture the characteristics of motion crowding (Bex & Dakin, 2005). Another study that examined relationship of the crowding effect with chromaticity found that when target and flankers had different chromatic properties, crowding effect is reduced (Kennedy & Whitaker, 2010). The authors indicated that this difference in the crowding is more than a color “pop-out” effect, as most of the other previous studies found (Kooi et al., 1994; Põder, 2007; Sayim, Westheimer & Herzog, 2008), rather, it suggests that processing of achromatic and chromatic information is segregated in the context of crowding. These studies shed light on the idea of multiplicity of crowding, as was explicitly demonstrated in a recent study by Greenwood and Parsons (2020) that crowding effect is not a singular process. Although Greenwood and Parsons (2020) did not address the relative sensitivity to crowding between color or motion, it seems to be the case in their results that the overall magnitude of color crowding was lower than motion crowding, consistent with the results of Experiment 2 in the current study.

In comparing the crowding effect across stimuli and tasks, it is important to match the performance levels across them. We note that the task difficulties were not perfectly matched across stimuli and tasks in our study despite our best attempt. We tried to mitigate this problem by including a practice session in the beginning of all tasks to ensure that all participants were able to perform above 80% accuracy for non-crowded conditions. In addition, we tried to reduce the effect of task difficulty by normalizing data across conditions. However, we acknowledge that the possibility that the pattern of results in our study was affected by differences in the task difficulty across the three conditions. Future studies with different stimuli and tasks than the ones used in the current study will help to address the generalizability of our conclusions.

In summary, results from Experiment 1 suggest that the processing in the M pathway may be more susceptible to spatial crowding effect compared to processing in the P pathway, this observation could be partially explained by the differential receptive field sizes between these two pathways at the same eccentricities. Using stimuli with different functional properties in Experiment 2, our results show that crowding was more severe in discriminating forms which requires integrating simple components into distinct items, compared to extracting motion direction and color information.

Chapter 4: Crowding and Visual Span in Dyslexia

Introduction

Dyslexia is a specific reading disability that has neurobiological origin and is characterized by difficulties with accurate and fluent word recognition, poor spelling and poor decoding skills, despite normal intelligence and age-appropriate education (American Psychiatric Association, 2013). Dyslexia is estimated to occur in 10% to 15% of the population (Shaywitz, Escobar, Shaywitz, Fletcher & Makuch, 1992), with some variations across different languages (Miles, 2004). Although the exact genetic pattern is still unknown, research on genetic linkage analyses have identified some genomic regions that are highly correlated with dyslexia (Francks, MacPhie, & Monaco, 2002). Some of the symptoms of dyslexia seem to be observed before children learn to read, supporting its neurobiological origins. Several longitudinal studies provided evidence that auditory-phonological skills and visual attentional processing abilities are impaired at the pre-reading stage in children who eventually develop dyslexia (Gallagher, Frith, & Snowling, 2000; Franceschini, Gori, Ruffino, Pedrolli & Facoetti, 2012; Carroll, Solity, & Shapiro, 2016). Given the longitudinal evidence, individuals' phonological skills and visual-attentional abilities are widely studied to better understand the characteristics of dyslexia. The current paper focuses on the visual processing difficulties associated with dyslexia.

It is also known that dyslexia frequently co-occurs with certain disorders such as ADHD (Germanò, Gagliano & Curatolo, 2010). Because of this comorbidity, reading problems are frequently misinterpreted as a result of other behavioral abnormalities. This confusion leads to the possibility of dyslexia going undiagnosed.

Diagnosing dyslexia is a complicated problem considering there is no physiology-based test that can provide an unambiguous quantitative diagnosis. Instead, a number of behavioral tests are used to identify dyslexia. However, reading is measured on a continuum and the cut-off point for abnormal reading skills is rather arbitrary (Siegel, 2006). Especially given the very large variability in the characteristics and severity of dyslexia, the distinction between people with dyslexia and typical readers vary across studies. While some studies use a subset of widely known standardized assessment tests such as Woodcock Johnson Assessments (Joanisse, Manis, Keating & Seidenberg, 2000; Visser, Boden & Giaschi, 2004; Joo, White, Strodtman & Yeatman, 2018), some other studies use a completely different battery of tests (Eden et al., 1996; Cestnick, 2001; Perrachione et al., 2016). Several research groups have also developed their own tests such as Castle and Coltheart Test (Castles et al., 2009) and York Adult Assessment Battery (Warmington, 2016), with the emphasis on specific skills based on their research questions. In addition to the variations in these tests, there are alterations in the assessments that are developed in other languages. Although vast majority of these assessments were verified for validity and test-retest reliability, the selection on what tests to use in a particular study and what cut-off point to apply to determine the dyslexia are not standardized, which inevitably generates inconsistencies among studies in terms of the definition of their dyslexia group.

In addition to arbitrarily selecting the assessments batteries, the dyslexia condition itself contains a great deal of variability. Dyslexia is a multilevel syndrome and experimental findings show that abnormalities exist at multiple levels and pathways (Galaburda, 1999). There is large variability in how dyslexia can be expressed in an individual relative to another, indicating dyslexia cannot be considered as a uniform condition. Over the last couple of decades, the subgroups of dyslexia have been discussed by many researchers and existence of distinct varieties have been shown (Castle & Coltheart, 1993; Lachmann, Berti, Kujala & Schroger, 2005; van Ermingen-Marbach, Grande, Pape-Neumann, Sass, & Heim, 2013). However, there is no consensus on what the exact subtypes of dyslexia are. Studies use different behavioral and neurological measures to address the heterogeneity in dyslexia and identify the subtypes (Williams, Stuart, Castles & McAnally, 2003; Lorusso, Cantiani & Molteni, 2014; Jednorog, Gawron, Marchewka, Heim & Grabowska, 2014). Moreover, the majority of dyslexia studies fail to acknowledge the varieties in their dyslexia group and attempt to draw a general conclusion from a small mixed sample which further contributes to inconsistencies in findings in the literature.

Due to all the variability in identifying the dyslexia group and the heterogeneity in the dyslexia condition itself, there is no consensus on the underlying mechanisms of dyslexia, despite the large number of studies. A recent review article discussed that there are more than eleven major theories on the mechanisms of dyslexia and each of them can be supported with several empirical studies (Ramus & Ahissar, 2012). In addition to different samples of subjects with different types of dyslexia, these discrepancies may also result from variations in the experimental paradigms.

Nonetheless, the majority of these studies emphasize that people with dyslexia predominantly suffer from two main information processing deficits: poor phonological abilities and impaired visual performance. The current paper aims to address particular characteristics of the latter: impaired visual performance. More specifically, this study investigates two aspects of visual performance relevant to reading: visual crowding and visual span in dyslexia.

Visual crowding, the inability to recognize objects in clutter (Whitney & Levi, 2011), is considered as the bottleneck of object recognition and predominantly occurs in the peripheral visual field. Although crowding can be observed with numerous different visual stimuli including letters, faces, moving versus static objects, with low or high spatial frequencies, in low or high contrasts, a prominent example of crowding's influence in daily life is in the domain of reading. As people make use of visual information in their periphery during reading (i.e., pre-viewing of upcoming letters while focusing on the current letters), crowding plays an important role in reading performance. Given this close relationship, crowding has been a topic of interest in the dyslexia research community and has become one of the most debated visual phenomena in this literature. Although there is empirical evidence for a stronger crowding effect in dyslexia (Bouma & Legein, 1975; Moores, Cassim & Talcott, 2011; Franceschini et al., 2012), several studies fail to find this difference (Shovman & Ahissar, 2006; Martelli, Di Filippo, Spinelli, & Zoccolotti, 2009; Doron, Manassi, Herzog, & Ahissar, 2015). Similarly, while some studies provided evidence for benefits of extra-large spacing for improving reading in dyslexia (Zorzi et al., 2012; Bellocchi, Massendari, Grainger & Ducrot, 2019), other studies showed that this improvement is not specific to the dyslexia population (Perea, Panadero, Moret-Tatay & Gómez, 2012; Hakvoort, van den Boer Leenaars, Bos & Tijms, 2017). A strong crowding effect is observed when objects are located close to each other in the periphery. Increasing the space between the objects reduces the “clutter” and helps object recognition. However, whether the same phenomenon applies to letter spacing and reading, and whether this effect is same or different for typical readers and people with dyslexia is still unclear. To better understand these inconsistent results, a small number of studies have attempted to address the differences in the experimental stimuli (Ziegler, Pech-Georgel, Dufau & Grainger, 2010) and variability in the dyslexia population (Joo et al., 2018). Joo and colleagues successfully identified a sub-group of participants with dyslexia who specifically suffered from elevated crowding and found that these individuals benefit from personalized interventions such as increased spacing between letters.

One of the main reasons why the perceptual system fails to resolve single letters in a crowded text is proposed to be the limited attentional resolution (He, Cavanagh & Intriligator, 1996). As visual spatial attentional abilities have been considered as a significant factor in dyslexia (Vidyasagar & Pammer, 2010; Franceschini et al., 2012; Stein, 2014), the link between crowding and dyslexia may help improve the understanding of how attentional mechanisms are associated with reading impairments. More importantly, identifying the sub-groups of dyslexia population in relation to their object recognition performance in crowding settings may play an important role on implementing training strategies to improve reading skills. For example, individuals showing elevated crowding may read faster when text is rendered with increased letter spacing. In this paper, our first goal is to investigate the differences in the crowding effect between individuals with and without dyslexia.

In addition to crowding, visual span is an important contributor to reading performance. Visual span is the number of text letters that can be recognized accurately on each eye fixation (Legge, Mansfield, & Chung, 2001). It has also been shown that the size of the visual span is closely related to reading speed (Legge et al., 2001; Yu, Cheung, Legge & Chung, 2007). Despite the close relationship between visual span and reading, surprisingly, visual span has never been investigated in the dyslexia population. On the other hand, visual attention (VA) span, which is the number of distinct visual elements that can be processed in parallel in a multi-element array (Bosse, Tainturier & Valdois, 2007), has been widely investigated. Multiple studies showed that individuals with dyslexia are impaired in their VA span skills (Bosse & Valdois, 2009; Lobier, Zoubrinetzky & Valdois, 2012; Chen, Schneps, Masyn & Thomson, 2016). Although visual span and VA span have some similarities, they are two different concepts addressing different mechanisms. Both measurements require individuals to fixate on a dot on the screen and verbally report the letters presented on the screen without moving their eyes. However, in visual span measurement, the letters are placed in close proximity as in typical reading text, allowing for crowding interference while in visual attention span, letters are presented with a distance in between them, minimizing the effect of crowding. Also, while reporting the letters, visual span requires the order to be correct while visual attention span is measured with no order constraint. Therefore, visual span measurement involves the effect of crowding and mislocation errors while visual attention span attempts to eliminate those errors and focus on the attentional abilities (for a detailed review, see Frey & Bosse, 2018). Our second goal is to study the visual span profiles in individuals with and without dyslexia. Considering the close nature of visual span and VA span, we expect to find similar findings; smaller visual span profiles in dyslexia.

In the current paper, we hypothesized that 1) the spatial extent of crowding in dyslexia would be greater than in typical readers, 2) the magnitude of crowding effect would generalize across different stimulus categories, suggesting that elevated crowding in dyslexia is general, not specific to language-related stimuli and 3) the visual span would be smaller in dyslexia, accounting in part for slower reading.

In addition to the three hypotheses stated above, we also aimed to investigate how the performance in these tasks may vary depending on the individuals. As mentioned, there are sub-types of dyslexia that may demonstrate differences in the visual task performance. These differences may result from various underlying forms of dyslexia or differences in the severity of dyslexia. However, as a result of unavoidable recruiting and scheduling constraints due to the Covid-19 pandemic, our data collection had to be suspended. With the current preliminary results, interpretation of individual differences in these visual tasks is not feasible. Nevertheless, preliminary data showed that the dyslexia population demonstrate overall impaired performance in all tasks compared to typical readers, with the caveat that the dyslexia group present considerable variability in their performance.

Methods

Participants

Twenty-one participants were recruited for this study. The participant pool consisted of two groups: dyslexia and control. Control group had ten participants (seven females), aged between 19-51 years (M=33.2), and were recruited from the University of Minnesota. Dyslexia group had eleven participants (six females), aged between 18-60 years (M=28.6), and all self-reported as having dyslexia. Participants in the dyslexia group were recruited through several channels such as through the University of Minnesota, the social media platform announcements and Learning Disabilities Association (LDA) Minnesota.

The comorbidity between dyslexia and ADHD is known to be high (Germanò, Gagliano & Curatolo, 2010). In dyslexia studies, it is the best practice to control for ADHD. In the current study, participants were asked to self-report whether they were diagnosed with ADHD, and six participants (five in dyslexia, one in control group) reported that they have ADHD. However, we could not exclude participants with ADHD due to the difficulties in recruiting subjects.

All twenty-one participants were assessed for reading and phonology abilities (assessments described later). Table 4-1 shows detailed information on the demographics and assessment scores of the participants.

All the participants were native English speakers with self-reported normal or corrected vision. They gave written informed consent to participate in accordance with the policies approved by the human subjects' review committee of the University of Minnesota.

	Control group (n=10)	Dyslexia Group (N=11)	t value
Age (years)	28.6 (3.59)	33.2 (3.94)	-0.86
Basic Reading Skills (percentile)			
Letter-Word Recognition	64.9 (6.8)	36.7 (7.2)	2.89**
Word Attack	59.9 (7)	23.1 (4.5)	4.45***
Reading Rate (percentile)			
Sentence Reading Fluency	75.2 (4.5)	33.3 (8.3)	4.42***
Word Reading Fluency	72.7 (4)	27.8 (7.2)	5.43***
Spelling (percentile)	59.1 (7.7)	34.8 (7.8)	2.21*
Phonological Skills (percentile)			
Word Discrimination	72.5 (2.5)	60.5 (6.7)	1.67
Phoneme Segmentation	48.9 (7.4)	19.8 (3.7)	3.51**
Phonological Blending	54.5 (9.9)	27.3 (8.6)	2.07
Castle and Coltheart Test (z-scores)			
Regular Word Reading	1.1 (0.2)	0.2 (0.4)	2.1
Irregular Word Reading	2.5 (0.1)	1.3 (0.3)	3.88**
Nonword Reading	-0.5 (0.1)	-1.2 (0.2)	3.69**

Table 4- 1: Characteristics of participants. Means and SDs of the two groups' scores in phonological and reading measures. Notes: *p <0.05, **p <0.01, ***p <0.001.

Stimuli, Apparatus and Procedure

The stimuli for the psychophysics experiments were presented using MATLAB with the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997) and displayed on a 21.5-inch iMAC (refresh rate: 75Hz, resolution: 4096 x 2304). The specifics of the stimuli were different for each experiment, which will be explained in the relevant sections. All stimuli were viewed binocularly from 40 cm in a dimly lit room. Viewing distance was maintained with the use of a chin rest. Six participants in the dyslexia group were tested at LDA and the rest of the participants were tested at the University of Minnesota Psychology Department. The exact same stimuli and apparatus were used to test in both locations.

Subjects participated in two experimental sessions: reading/phonology assessments and psychophysics measurements, which consisted of three experiments. Participants were either tested in two separate days for two experimental sessions, or they were tested in the same day with at least an hour break in between the sessions.

During the reading/phonology assessments session, participants completed a battery of phonology, reading and spelling tests, which were subsets of Woodcock Johnson-IV (WJ-IV) Tests of Achievement, Test of Auditory Processing Skills (TAPS-3) and Castle and Coltheart Test 2 (C&C2). In the psychophysics measurements, participants were tested in three experiments: crowding, visual span and Flashcard reading test. The order of the psychophysics tests was counterbalanced across participants. The details of both cognitive assessments and psychophysics experiments are described in the next sections.

Assessments

Participants were assessed in nine different tests to measure their skills in auditory processing, phonology, reading and spelling tasks. The order in which the tasks were administered was the same across all participants.

Letter-Word Identification

This test is taken from WJ-IV Test of Achievement battery and addresses Basic Reading Skills, Orthographic Awareness and Phoneme-Grapheme Knowledge categories. It is designed to measure the level of skills in reading regular words. In this test, individuals read aloud the words that are listed on a page. The words become gradually harder in terms of their pronunciation (e.g., easy pronunciation: social, difficult pronunciation: impetuosity). Real-word reading performance is measured, highlighting the typical reading performance including both phonological and semantics abilities.

Word Attack

This test is taken from the WJ-IV Test of Achievement battery and addresses Basic Reading Skills, Orthographic Awareness and Phoneme-Grapheme Knowledge categories. It is designed to measure the level of pseudo-word (non-words that conform to English spelling patterns) reading skills. In this test, similar to the previous test, individuals read aloud the words that are listed on a page. This time the words are pseudowords (nonwords that comply with English language rules). The words become gradually harder in terms of their pronunciation (e.g., easy pronunciation: floxy, difficult pronunciation: pretrationistic). Pseudoword reading performance is measured, highlighting the phonological knowledge without the help of semantics.

Sentence Reading Fluency

This test is taken from the WJ-IV Test of Achievement battery and addresses Reading Fluency and Reading Rate categories. It is designed to measure the level of skills in reading accurately with sufficient speed. This test is timed. In this test, individuals silently read sentences listed on a page, and decide if the statement is correct or not. It is aimed to measure both accuracy and speed.

<u>Word Reading Fluency</u>

This test is taken from the WJ-IV Test of Achievement battery and addresses Reading Rate category. It is designed to measure the level of skills in reading accurately with sufficient speed. This test is also timed. In this test, individuals silently read four words in each line, many lines listed on a page, and decide which pair of those four words are semantically related to each other and cross those words off (e.g., words in a line: “two”, “green”, “red”, “cat” with the expectation of “green” and “red” should be crossed off). It is aimed to measure both accuracy and speed.

<u>Spelling</u>

This test is taken from WJ-IV Test of Achievement battery and addresses Spelling and Orthographic Awareness areas. It is designed to measure the level of spelling skills. In this test, individuals listen to a voice on a recording pronouncing words, gradually increasing in difficulty to spell (e.g., easy spelling: actually, difficult spelling: soliloquy)

<u>Word Discrimination</u>

This test is taken from the Test of Auditory Processing Skills (TAPS-3). It is designed to assess the level of ability in discerning phonological differences and similarities within word pairs. In this test, individuals listen to a recording of pronunciation of many pairs of words and decide whether the two words are the same or different (e.g., same: decision-decision, different: confirm-conform)

<u>Phonological Segmentation</u>

This test is taken from the Test of Auditory Processing Skills (TAPS-3). It assesses how well individuals can manipulate phonemes within words. In this test, individuals are asked to pronounce a word and then delete a part of that word (e.g. say “finish” and now say it without “ish”).

<u>Phonological Blending</u>

This test is taken from Test of Auditory Processing Skills (TAPS-3). It assesses how well individuals can synthesize a word given the individual phonemic sounds. In this test, individuals hear a set of sounds and are expected to say the word that those sounds make (e.g., sounds “sh / i / p / m / e / n / t”)

<u>Castle and Coltheart Test 2</u>

This test was developed by Anne Castle and Max Coltheart in Macquarie University. It is designed to directly assess how well individuals can use the lexical and the non-lexical procedures. This test includes 40 each of regular words, irregular words and pseudo-words. Individuals read a list of words printed on paper, gradually increasing in difficulty in each category. For example, for the real word category, easy: free, difficult: quaver; for the irregular word category, easy: blood, difficult: chassis; for the pseudo-word category, easy: roft, difficult: spoltchurb. Individuals are scored based on whether they pronounce the words correctly or not.

Table 4-1 shows the group averages on the general assessment scores for both dyslexia and control groups (the detailed information about the individual scores in each assessment can be found in Appendix 6). In Table 4-1, Castle and Coltheart Test scores are shown in z-scores and the rest of the assessments scores are shown in their percentile ranks.

Flashcard Reading Speed Test

The flashcard reading paradigm was used to test participants' reading speed. In this test, sentences were presented on the screen for a certain amount of time and participants were asked to read the sentence aloud. After the sentence disappeared from the screen, participants were allowed to complete their verbalization. When participants completed their verbalization, the experimenter noted the number of correct words read for that sentence. Participants clicked the mouse to initiate the next trial.

The sentences were selected from MNREAD sentence pool (Mansfield & Legge, 2007) and were presented in Times New Roman font with an x-height of 1º. These sentences included 60 characters and an average of 10 words. Each sentence was rendered on the screen in three lines and justified configuration (Figure 4-1). None of the sentences was presented twice for a given participant. Participants were allowed to make eye movements during reading.

Each participant completed a practice run, which included 12 practice sentences. In the main experiment, participants read a total of 36 sentences. Each sentence was presented on the screen in one of the six predefined duration times. Each presentation duration was repeated three times. The order in which participants were presented different duration times was randomized across participants. The sentence and the duration time pairs were not fixed, in other words, while sentence#1 may be presented for 0.8 seconds to one participant, the same sentence could be presented for 3 seconds to another participant.

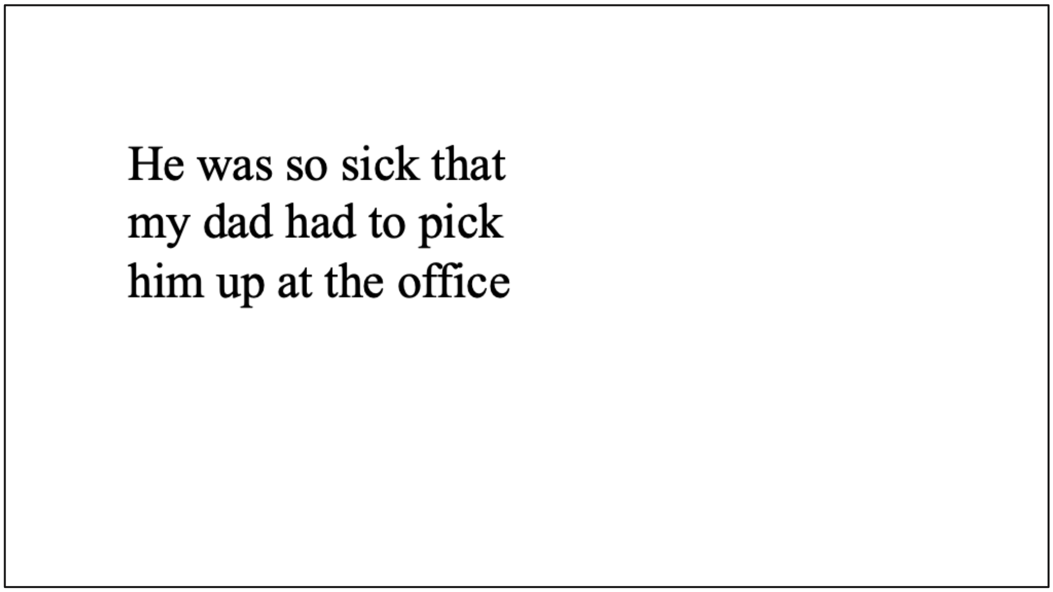


Figure 4- 1: Flashcard reading test- sample stimuli

The presentation durations were usually 0.4, 0.8, 1.2, 1.6, 2.4 and 3 seconds. Each participant was expected to be able to read no more than 30% correct at the shortest duration and at least 80% correct at the longest duration. For a small number of participants, the range of the duration times was adjusted to make sure they would be able to read the full sentence in the longest duration.

The data were fit with a psychometric function to obtain the maximum reading speed for each participant. More details on this calculation are given in the Data Analysis section.

Crowding Experiment

The crowding experiment measured the recognition performance in the peripheral visual field. In this experiment, three different stimulus categories were used: letter, unique symbols and square-wave gratings (Figure 4-2). Letter stimuli consisted of 6 uppercase letters (H, K, P, F, X, Z) in DejaVu Sans Mono font. Symbol stimuli consisted of 6 unique symbols which were created by Castet and colleagues with a one-to-one matching perimetric complexity between the letters and symbols (Castet, Descamps, Denis-Noël & Cole, 2017). Grating stimuli were generated using MATLAB and they were square-wave, high contrast gratings with the spatial frequency of 0.5 cycles per degree. The grating stimulus may take one of the six different orientations (0º, 90º, 20º, 70º, 340º, 290º). All stimuli, regardless of their category, were 2 visual degrees in size and presented on a gray background.

The target stimulus was presented in the periphery at 11º, horizontally aligned with the fixation point, either on the right or left side of the fixation in each trial. Two distractor objects, also called flankers, were placed on the right and left of the target stimulus (Figure 4-3a). Target and flankers were randomly chosen from the six possible options within their category. Within each trial, no target and flankers were identical to each other.

The experiment consisted of three sessions, with each session only including one type of stimuli. The order of the sessions was counterbalanced across participants. Each session consisted of 5 blocks, varying in their inter-stimulus spacing (the distance between the target and flankers). One of the blocks was defined as the baseline, as there was no flanker presented. In addition to the baseline block, the spacing in the other

4 blocks were 2º, 4º, 6º and 8º. Each block consisted of 40 trials, with a total of 200 trials for each condition.

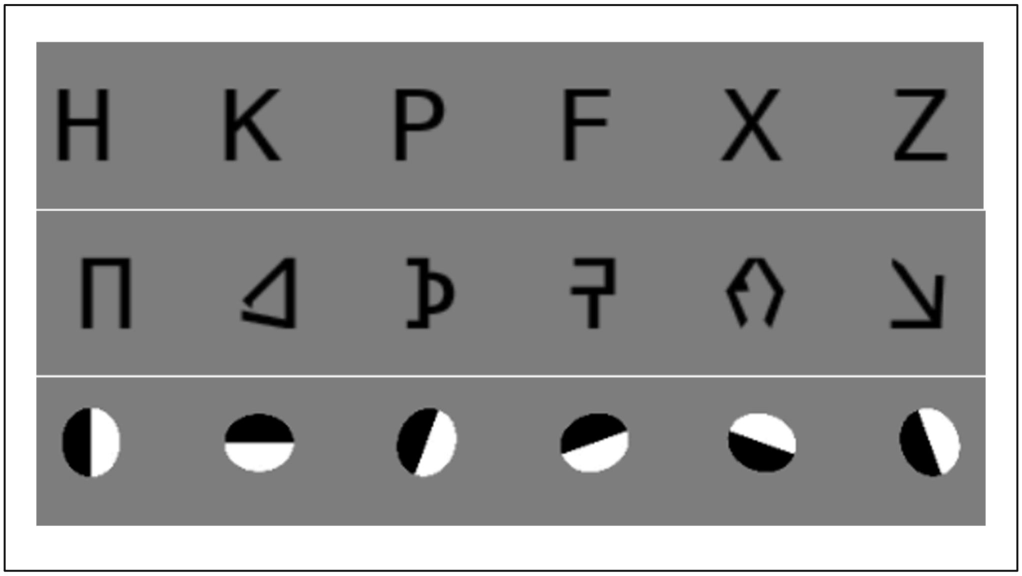


Figure 4- 2: Three stimulus categories for the crowding experiment.

The experiment was completed with binocular viewing. Participants were instructed to fixate at the black dot (visual angle of 1º) in middle of the screen throughout the whole experiment. While fixating at the middle, participants were asked to identify the target presented in the periphery. Target and flankers were presented on the screen for 100ms. After the presentation in the periphery, the response screen was displayed. In this screen, four options were given to the participants, with one of them being the correct response (Figure 4-3b). Participants were allowed to move their eyes and look at the options during the response period. They were instructed to give their response by using the keyboard within 2 seconds. The response keys were located in the keyboard spatially representing the response options on the screen. Immediately after the response, the next trial was presented.

a)

b)

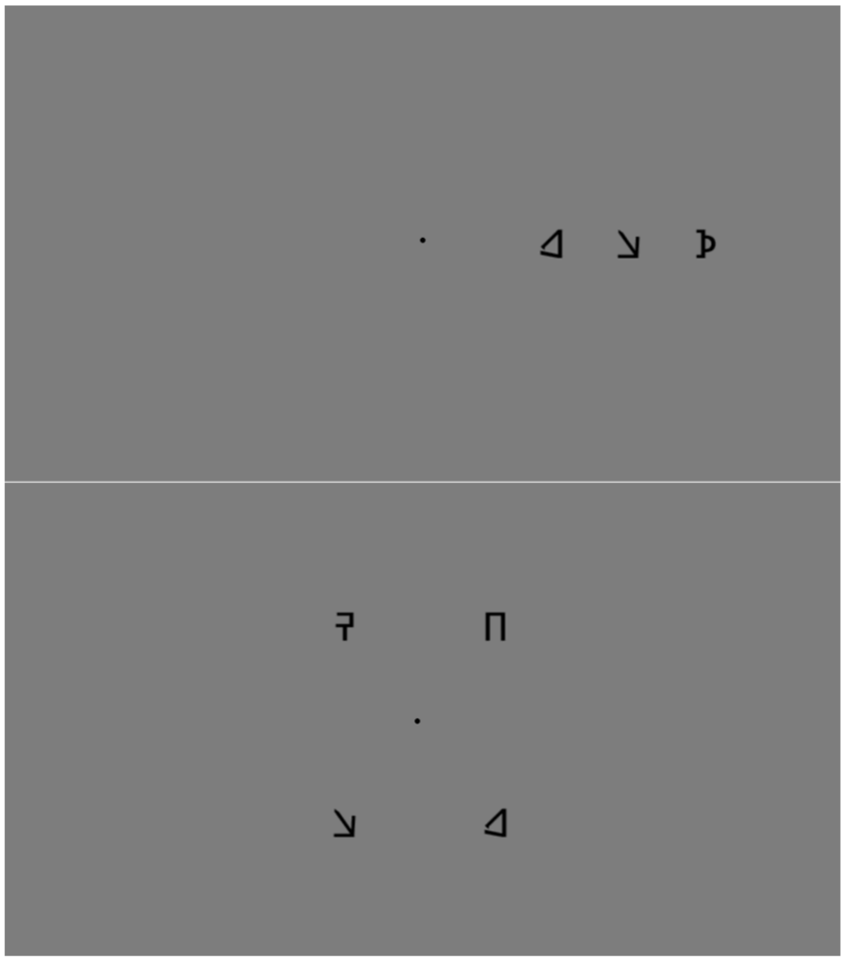


Figure 4- 3: The demonstration of a) an experimental trial and b) the response screen in the crowding experiment symbol condition

Before the main experiment, participants completed a practice run for each of the three stimuli categories to get familiar with the stimuli and the task. Practice run consisted of four blocks including the baseline block, 4º, 6º and 8º inter-stimulus spacing blocks. Each block consisted of 20 trials. Participants, who reached 80% accuracy in the baseline block for each stimulus category, continued to the main experiment. Participants who did not reach to 80% accuracy repeated the practice run. While majority of the participants were able to reach 80% accuracy at the first run, a few participants repeated the practice run two or three times. All participants reached 80% accuracy in less than four runs of the practice.

The accuracy scores were recorded by the program for each block and each stimulus category.

Visual Span Experiment

The visual span experiment measured the visual span size of participants in three different conditions: single letter recognition, trigrams with full report and trigrams with partial report. The stimuli used in all the conditions consisted of black lowercase letters on a white background. The letters were rendered in Courier font. Letter spacing was 1.16 * x-width (standard spacing for Courier).

Similar to the crowding experiment, participants were asked to fixate at the middle of the screen, on a gap in between the two dots. Letter stimuli were presented in one of the possible 11 slots, which were horizontally aligned with the fixation gap, as shown in Figure 4-4a. The slot on the fixation midline is labeled 0, and left and right slots to the fixation are labeled with negative and positive numbers, respectively. Letters were presented either in isolation (single letter recognition condition) or in triplets (trigrams with full or partial report). Letters were randomly chosen from 26 possible English lowercase letters and were presented for 100ms.

Participants clicked the mouse to initiate the trial. After the beep sound, the letter stimuli were presented. After the stimuli presentation, participants verbally reported the letters to the experimenter. No feedback was given. For the next trial, participants clicked the mouse again.

In the single letter condition, participants reported the single letter presented on the screen. In the trigram full report condition, participants reported all three letters. Participants were given correct scores as long as both the letter identity and the location in the trigram were correct. In the trigram partial report condition, participants were only asked to report the middle letter of the trigram. Each condition consisted of 77 trials, with 7 repetitions on each slot from -5 to 5, including 0 (the midline).

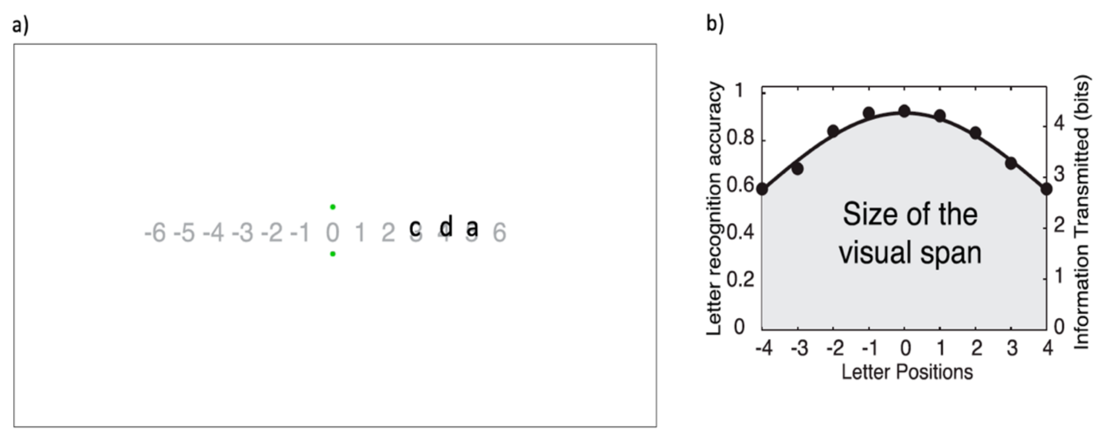


Figure 4- 4: a) The demonstration of a trial in visual span experiment when the stimuli are presented at letter position 4. The gray numbers indicating the letter positions are shown here for demonstration purposes, the participants in the experiment were not presented with these numbers. b) diagram of the visual span. A visual span profile is plotted as letter-recognition accuracy versus letter position and then converted to information transmitted (bits).The area under the curve is the size of the visual span. (Adapted from He et al., 2013.)

Letter recognition accuracy was calculated in each condition. Percent correct for the visual span in each letter position was obtained by summing across all trials as the letters at a given position are sometimes inner, sometimes middle and sometimes outer letters of the trigram). Proportion correct was plotted as a function of letter positions to obtain the visual span profiles for each participant (for an example, see Figure 4-4b).

Letter recognition accuracy was also converted to information transmitted in bits using letter-confusion matrices measured by Beckmann (1998). The chance level performance (one out of 26 = 3.8% accuracy) corresponds to 0 bits of information and 100% accuracy corresponds to about 4.7 bits.

The transformation of proportion correct letter recognition to bits of information conversion was done using the following formula:

Information transmitted in bits = -0.036996 + 4.6761 × letter recognition accuracy

The visual span size was quantified by summing across the information transmitted by the 11 slots of the profile.

Data Analysis

The analyses on the crowding experiment, visual span experiment and flashcard reading speed test were run separately.

In the crowding experiment, proportion correct was calculated for each inter-stimulus spacing block in all three stimulus categories. The accuracy data were also fitted as a function of target-flanker spacing with a cumulative Gaussian sigmoid curve using the Psignifit toolbox software for MATLAB (Wichmann & Hill, 2001) to obtain the critical spacing value for each participant in each category. Critical spacing is defined as the minimum required distance between the target and flankers to reach criterion percent correct. In this experiment, the critical spacing was calculated with the criterion of 80% correct. A linear mixed effect model was performed to see the difference in critical spacing between the two groups and between the stimuli categories. Group and stimuli categories were defined as fixed effects, while the participants were included in the model as random effects.

In the visual span experiment, letter recognition accuracy was calculated for all three conditions in both proportion-correct and information transmitted in bits. Visual span size was also determined by summing across the information transmitted by the 11 slots of the profile. One-way ANOVA was conducted to examine the difference between the two groups in their visual span sizes. The impact of the letter position within the trigram on visual span was also analyzed. At a given trigram position, trigrams included three spots: the center letter of the trigram (middle), the nearest spot to the midline (inner) and the farthest spot from the midline (outer). The visual span sizes were calculated, and visual span profiles were obtained for each letter position within the trigram.

Additionally, a decomposition analysis was conducted to determine the factors that account for the difference between perfect performance and the actual performance (visual span profiles) in the letter recognition task. Through this decomposition analysis, the impact of three factors on visual span profiles were investigated: acuity, mislocation errors and crowding. Acuity is measured through the errors in the single letter condition. Mislocation errors are defined as reporting the correct identity of an adjacent letter instead of the target letter. Lastly, crowding factor in visual span is assessed as incorrect identifications after taking into account mislocation errors and acuity errors (for a detailed description of the analysis, see He, Legge & Yu [2013]).

In the flashcard reading measurement, the number of correctly read words were plotted as a function of reading time by fitting a psychometric function. Maximum reading speed was determined at exposure time that yielded 80% of the words read correctly. Reading speed was computed according to the following formula:

Reading speed (in words per minute):
60 / criterion exposure time for word (in seconds)

The average maximum speeds for both dyslexia and control groups were calculated.

The reading and phonology assessments were scored based on the specific instructions of the tests. The scores for the subsets of Woodcock Johnson IV test and Test of Auditory Processing Skills 3 (TAPS-3) were converted to standardized scores. The z-scores were calculated for the Castle and Coltheart Test.

Results

Assessments

As Table 4-1 demonstrates, participants with dyslexia overall showed poorer performance compared to typical readers in all assessments. Specifically, in a number of measurements (i.e. Letter-word recognition, Word Attack, Sentence reading fluency, word reading fluency, phoneme segmentation, CC2 irregular word reading and CC2 nonword reading), average percentile rank of participants with dyslexia is significantly lower than average in typical readers ($p < 0.01$).

Both dyslexia and control groups showed variabilities within their own group. Especially in the dyslexia group, participants showed considerable variance in many assessment categories. Figure 4-5 demonstrates the individual differences in the Woodcock Johnson IV assessment categories for both dyslexia (left panel) and control (right panel) groups. These variabilities support our hypothesis that dyslexia is a complex syndrome in which individuals with dyslexia show different characteristics, possibly suggesting some sub-types of the condition. However, due to our small number of participants, our study does not have enough power to clearly demonstrate these sub-types.

As it can be seen in Figure 4-5, there are a few participants in the dyslexia group who score higher in some assessments than a couple of participants in the control

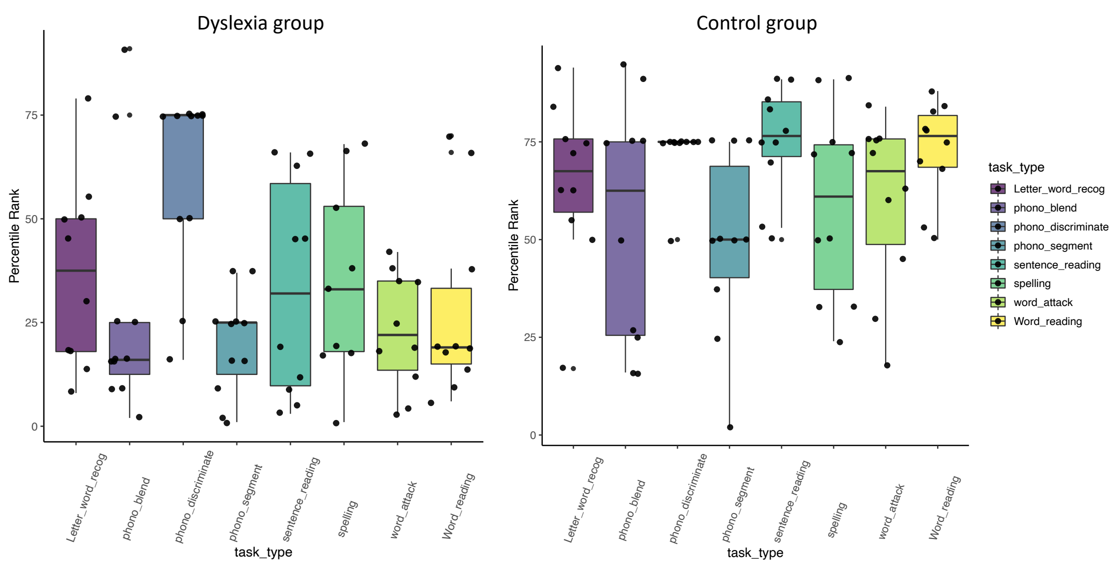


Figure 4- 5: Scores from the Woodcock Johnson IV, Tests of Achievements. Left panel shows the dyslexia group including individual data points and the group average. Right panel shows the control group individual data points and the group average.

group. In the current study, participants were assigned to the dyslexia and control groups based on their self-report. In the scenario where some of these assessments are chosen as criteria to decide as having dyslexia or not, the division of the participants might have been slightly different. Nevertheless, the results reported in this chapter are based on the two groups that were assigned through self-report. The individual scores in each assessment, along with the reading speed measured by the Flashcard Reading Speed Test, can be found in Appendix 6.

Some of the scores obtained through these assessments are highly correlated with each other, as well as the results from some of the psychophysics experiments. Figure 4-6 shows an extensive correlation heatmap.

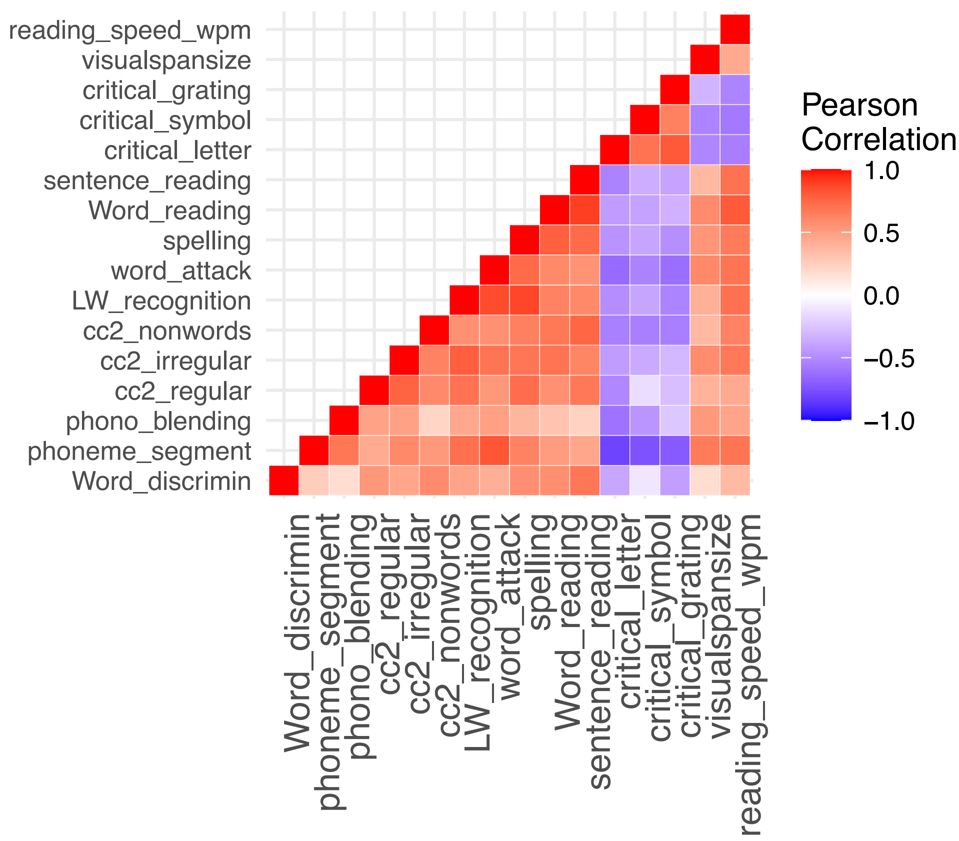


Figure 4- 6: Correlation heat map including all the standardized assessment scores and the values from psychophysical experiments

Flashcard Reading speed

As expected, the average reading speed measured by the Flashcard test was significantly slower for the dyslexia group compared to control group ($F(1,19)=36.916$, $p=0.000$). Figure 4-7 shows maximum reading speeds for individual participants along with the group average for both dyslexia and control groups.

Reading speed measures obtained by the Flashcard test were highly correlated with the reading assessments (i.e. Word Reading Fluency, Sentence Reading Fluency, Letter-Word Recognition, Word Attack and Spelling measurements from the Woodcock Johnson IV battery and Irregular and Nonword Reading Tests from Castle and Coltheart Test) and phonological assessments (i.e. Phonological Segmentation test from Test of Auditory Processing Skills battery). These correlations are shown in Figure 4-6. Individual scores on the various assessments and the Flashcard reading speed test are listed in Appendix 2.

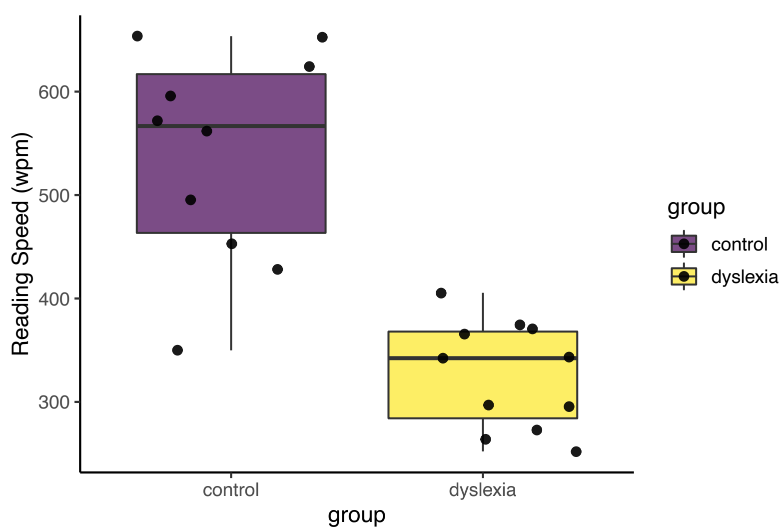


Figure 4- 7: Reading speeds measured with the Flashcard reading test including individual data points and group averages.

Crowding

Figure 4-8 demonstrates the main results of the crowding experiment. The recognition accuracy in the dyslexia group is overall worse than the control group in all of the stimulus categories.

A linear mixed effect model was performed to see the difference in critical spacing between the two groups for all stimulus categories. As seen in the Figure 4-9, the average critical spacing in the dyslexia group was significantly greater compared to the control group in both letter ($F(1,19)= 4.981$, $p= 0.038$) and symbol ($F(1,19)=5.085$, $p= 0.036$) conditions, but not in the grating condition ($F(1,19)= 2.361$, $p= 0.141$).

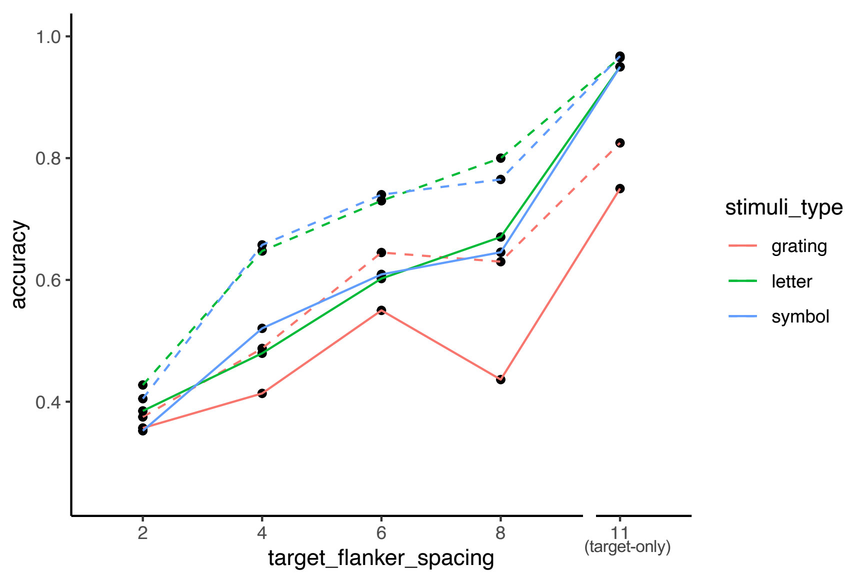


Figure 4- 8: Results of the crowding experiment. Accuracy performeance is plotted as a function of target flanker spacing. Dashed lines represent the control group, solid lines represent the dyslexia group.

On the other hand, we did not find any difference in the recognition performance between the letter and symbol categories within the same group for both control and dyslexia groups. In both groups, the performance for the grating condition was significantly lower compared to letter and symbol conditions (*p*=0.001). In the grating condition, the recognition performance in both groups had an unexpected decrease in the largest inter-stimulus spacing. This pattern will be discussed more in detail in the Discussion section.

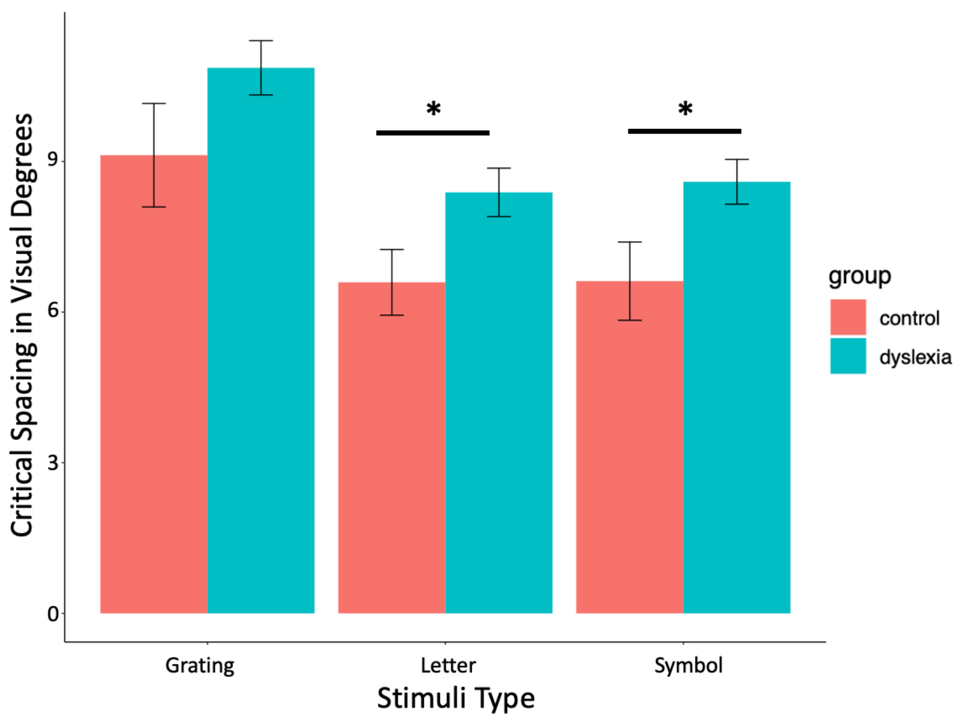


Figure 4- 9: The difference between the dyslexia and control group in average critical spacing values for all three stimulus categories.

While there is a considerable variation in individual performance within both groups, the participants in the dyslexia group tend to have lower performance compared to the control group. Individual critical spacing values are shown in Appendix 3 for both groups in all stimulus categories.

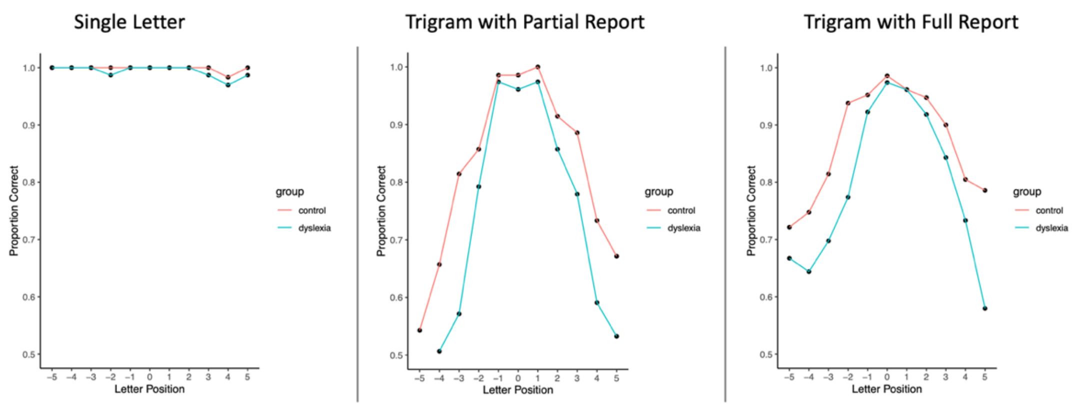
Average critical spacing values for both groups in all conditions were correlated with a number of phonological and reading assessments. Average critical spacings in the letter and symbol conditions were also correlated with visual span size ($r(19) = -0.49$, $p = 0.02$, $r(19) = -0.48$, $p = 0.03$, respectively). Appendix 4 demonstrates multiple correlations among all psychophysics measurements and all phonological/reading assessments.

Reduced Visual Span Size in Dyslexia

Figure 4-10a shows the average visual span profiles for the control and dyslexia groups tested with both single letters and trigrams including partial and full reports. The trigram visual span profiles in the dyslexia group are narrower compared to the control group, although we do not see a difference in single letter recognition performance between the two groups. Individual visual span profiles of dyslexia subjects compared to the control group average can be seen in Figure 4-10b. Despite the individual differences, all participants in the dyslexia groups show a narrower visual span profile compared to the average profile of the control group.

One-way ANOVA test was performed on visual span sizes for single letters and trigrams. We found a significant main effect of group ($F(1, 19)=7.11$, $p=0.015$) in the trigrams while there was no significant difference between the groups for the single letter measures. Figure 4-11 demonstrates the comparison between the two groups for the visual span size measurements for single letters and trigrams with both partial and full reports. The visual span profiles were also analyzed when tested with the partial report paradigm. As in full-report trigram visual span measures, one-way ANOVA test showed that there is a significant difference between dyslexia and control group visual span size ($F(1,19)=8.85$, $p=0.008$) when participants were asked to only report the middle letter of the trigram.

a)



b)

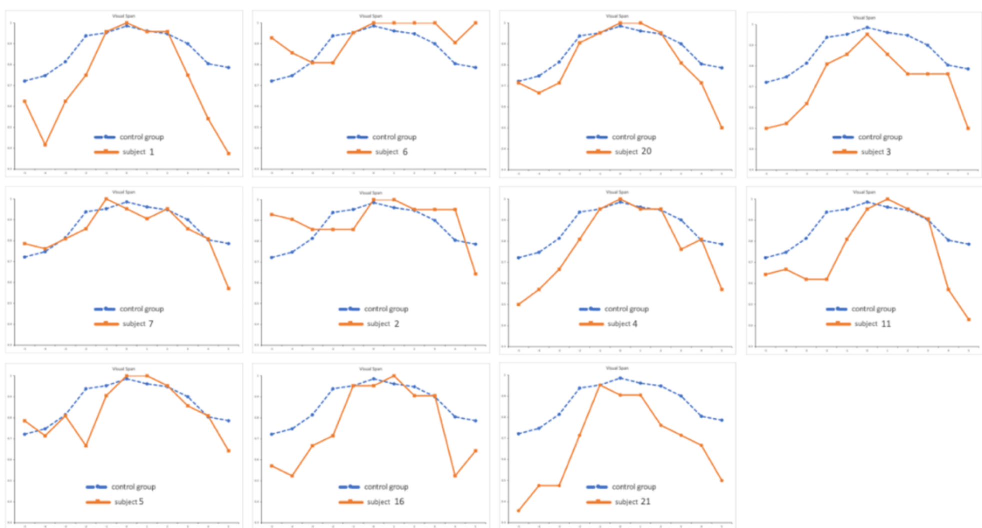


Figure 4- 10: a) Visual span profiles (letter-recognition accuracy as a function of letter positions) for single letters (the orange lines) and trigrams (the blue lines). The dotted lines represent the control group while the solid lines represent the dyslexia group. b) Individual visual span profiles for participants in the dyslexia group compared with the average visual span profile of the control group.

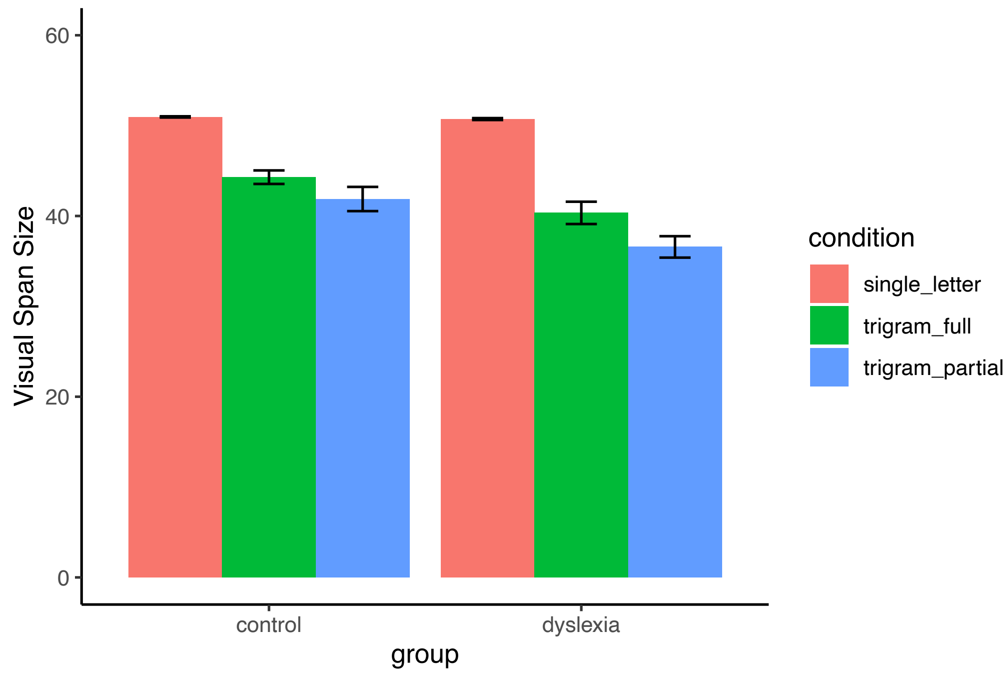


Figure 4- 11: Comparison between the visual span size when tested with trigrams as well as single letters for the control and dyslexia groups.

Figure 4-12a shows the separate visual span profiles, averaged within groups, for inner, middle and outer letters. In both dyslexia and control groups, highest accuracy (broadest profile) is observed for the outer letters and the lowest accuracy is observed for middle letters. This pattern of visual span profiles in relation to the letter positions within the trigram is consistent with the previous studies (Legge, Mansfield and Chung, 2001). We see a similar pattern for inner/middle/outer letter visual profiles between the dyslexia and control groups, with visual span profiles being consistently narrower for the dyslexia group. Interestingly, the magnitude of the difference between the control group and dyslexia group is not the same across different letter positions within a trigram. As Figure 4-12b demonstrates, the difference between the two groups is significant in the inner (p=0.006) and middle (p=0.04) letter positions, while it is negligible in the outer letter position (p>0.05).

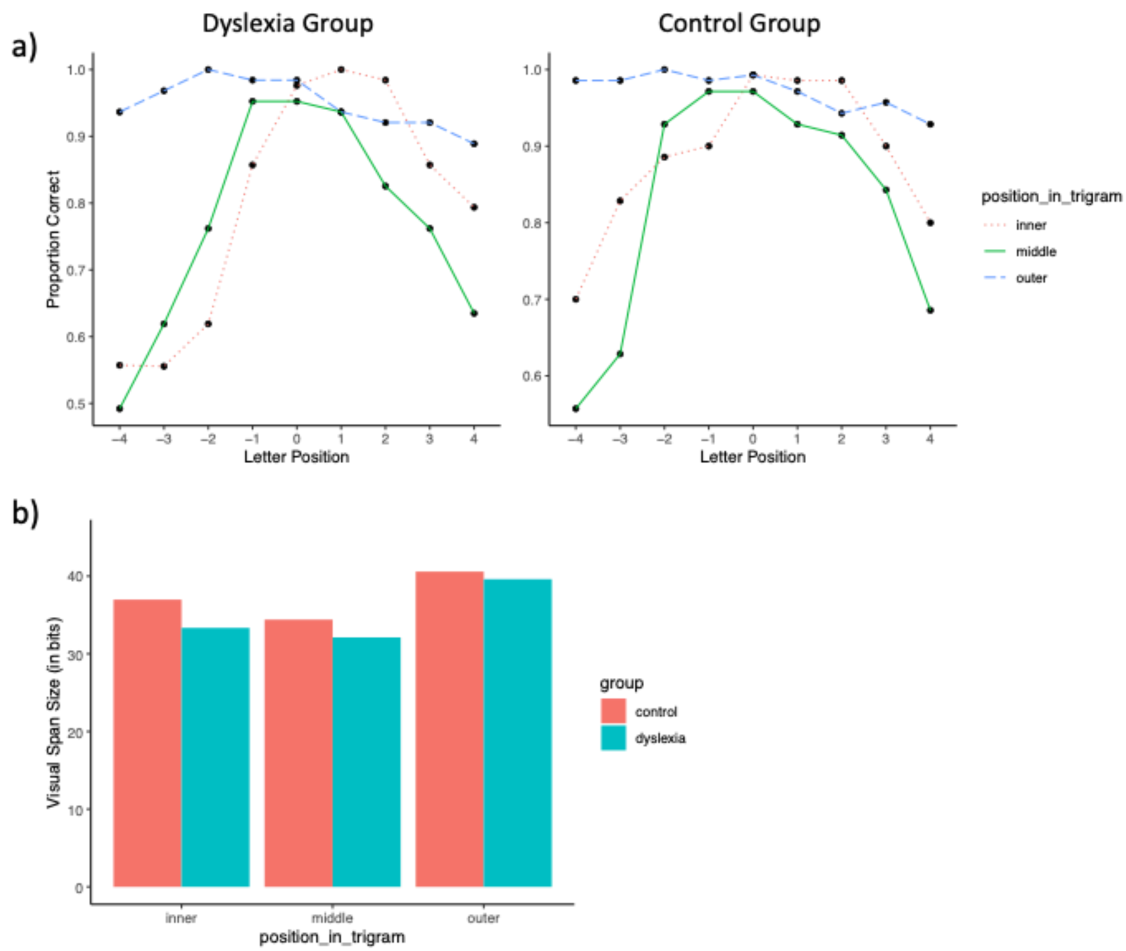


Figure 4- 12: a) Effect of letter position within the trigram: Separate visual span profiles are shown for inner (dotted), middle (solid) and outer (dashed) letters in the both dyslexia and control groups. b) The comparison between the average visual span sizes of dyslexia and control groups for inner/middle/outer letter positions of the trigram.

A decomposition analysis was performed on visual span profiles to assess the contributions of three factors: acuity, mislocations and crowding. This analysis was adapted from He et al. (2013), aiming to examine what factors determine participants' visual span sizes. Consistent with previous results, for both dyslexia and control groups, the contribution of crowding (control: 5.93 bits, dyslexia: 8.60 bits) was larger than the contributions from mislocations (control: 0.72 bits, dyslexia: 1.42 bits) and acuity (control: 0.08 bits, dyslexia: 0.67 bits). Although in both groups the crowding effect is the largest, the contribution of mislocations and acuity is higher in dyslexia group compared to the control group, however this difference between the groups is not statistically significant.

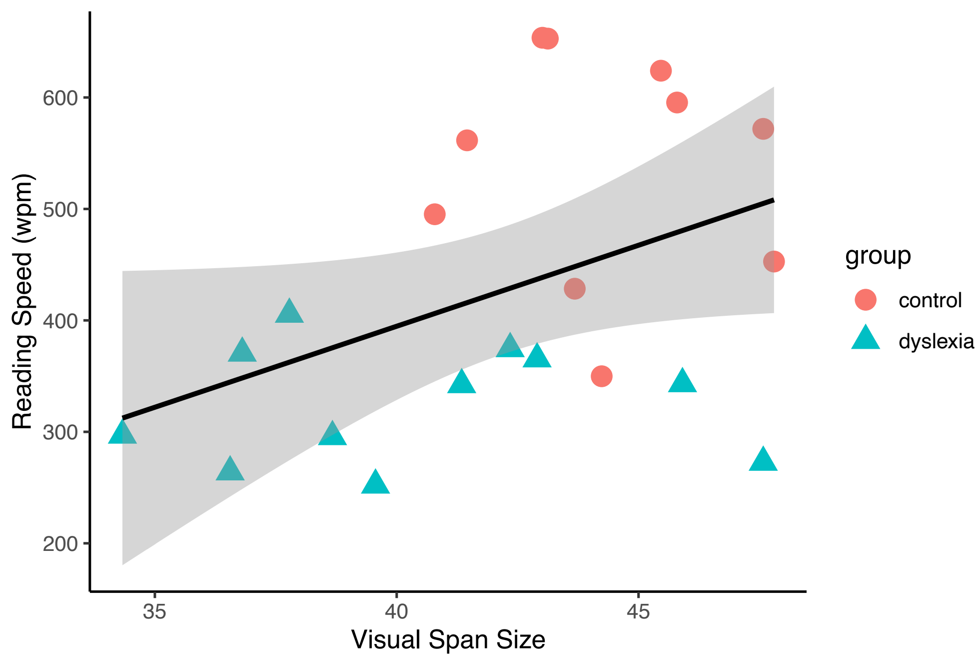


Figure 4- 13: Correlation between the reading speed (wpm) and visual span size for all the participants.

Consistent with the literature, visual span sizes in our study are found to be correlated with flashcard reading speeds ($r(19) = .42$, $p = .058$) (Figure 4-13). This correlation is observed to be stronger when it is calculated with all of the subjects, as the correlation coefficient value decreases when the correlation is separately computed within two groups.

Lastly, similar to critical spacing values, visual span sizes in both groups were also correlated with a number of phonological and reading assessments. Correlation coefficients between psychophysics measurements and phonology/reading assessments are provided in Appendix 4.

Summary of Results

We tested participants in dyslexia and control groups in a battery of reading/phonology assessments and three different psychophysics tests: 1) flashcard reading test, 2) crowding and 3) visual span. As summarized from the above analyses, we found that:

- The reading/phonology assessments showed that the dyslexia group had overall lower scores compared to the control group; however, large variability was observed within both groups.
- Participants with dyslexia showed slower reading performance in the Flashcard reading test compared to the control group, confirming the assignment of subjects to two groups.
- Participants with dyslexia show worse performance in object recognition in crowding conditions. In letter and symbol stimulus categories, critical spacing of participants in the dyslexia group were significantly larger than participants in the control group, indicating that dyslexia group needed larger spacing between target and flankers to achieve a criterion level of recognition. Regarding the

differences between stimulus categories within groups, both groups showed similar patterns such that the recognition performance in letter and symbol conditions was not different than each other, while grating condition yielded the worst performance for both groups. within each group. These findings support our hypotheses 1) the spatial extent of crowding in dyslexia is greater than in typical readers and 2) the magnitude of crowding effect generalizes across different stimulus categories, suggesting that enhanced crowding in dyslexia is general, not specific to language-related stimuli.

- Participants with dyslexia have narrower visual span profiles compared to the participants in the control group when tested with trigrams. Although participants did not differ in their single letter recognition performance, people with dyslexia were significantly worse at recognizing the letters when presented in triplets. This suggests that visual span, which is highly correlated with reading performance, is narrower in dyslexia due to crowding and mislocation errors, not simply due to acuity. This finding supports our third hypothesis: the visual span would be smaller in dyslexia, accounting in part for slower reading.

Discussion

Crowding Effect in Dyslexia

In the crowding experiment, three different stimulus categories were used: letter, unique symbols with matched perimetric complexity to letters and square-wave gratings. Our results showed that participants with dyslexia on average had worse recognition performance and higher critical spacings for letters and symbols compared to the typical readers. The difference between the groups are small, although statistically significant. These results are consistent with the previous studies that found difference in the crowding performance between people with dyslexia and typical readers (Bouma & Legein, 1975; Moores, Cassim & Talcott, 2011; Franceschini et al., 2012). Importantly, our results for the grating condition are also in line with the studies in which no difference was found in the recognition performance in crowding settings between the dyslexia and control groups (Shovman & Ahissar, 2006; Doron et al., 2015). Our results indicate that the difference between the dyslexia group and typical readers in crowding effect partly depends on the type of the stimulus.

It is important to note the different pattern of results for different stimulus categories in our study. Processing letters or symbols used in our experiment requires feature conjunction for a complete recognition while gratings only include single information which is the orientation that needs to be processed. In line with Martelli et al. (2009), these findings provide evidence for the idea that people with dyslexia may suffer from abnormalities in integrating information in the periphery.

Furthermore, the grating condition yielded the lowest performance in both groups allowing us to avoid the ceiling effect in this orientation discrimination task, which is supposedly less complex than a full object identification task, to be able to see the difference in the performance, if any. In the grating condition, participants' performance decreased in the largest inter-stimulus spacing, which was surprising. One possible interpretation of this result is that participants were distracted by the inner flanker, which is the one closer to middle of the screen, as it was right next to the fixation where participants were looking at. In this large spacing condition with a continuous stimulus (the difference among the six orientations of the grating is on a continuum, as opposed to categorical stimulus), like gratings in this case, the flankers may impair the performance not by crowding the target stimulus but by distracting the observer through being in their foveal radius. However, more studies may be needed to see this discrepancy on the effect of inner flanker between the stimulus categories.

To sum up, our results suggest that people with dyslexia have intact perceptual skills in the earlier stages of information processing, while they may have impaired abilities in the later stages which usually requires feature integration and larger involvement of attentional mechanisms. A somewhat different interpretation may be that, since symbols and letters are formed by feature conjunctions, their processing requires focused attention, in contrast to gratings. Therefore, the finding that participants with dyslexia had more severe crowding with symbols and letters may suggest a potential attentional deficit.

Visual Span Profiles in Dyslexia

Visual span has been shown to be correlated with reading speed (Legge, Mansfield, & Chung, 2001; Yu, Cheung, Legge & Chung, 2007) when studied with people with no known reading disabilities. To our knowledge, for the first time in the literature, we examined the visual span profiles in a dyslexia group. Our results showed that individuals with dyslexia have narrower visual span profiles on average compared to the typical readers.

In our visual span experiment, we measured participants' performance in three different ways: single letter recognition, reporting only the middle letter of a trigram and reporting all three letters in a trigram. There was no difference in single letter recognition between the two groups, indicating that dyslexia group does not simply have impaired letter recognition which can be associated with acuity in peripheral vision. Participants with dyslexia showed, however, poorer performance in both partial and full report of the trigram recognition, yielding narrower visual span profiles. It is important to note that participants had to report only one letter in the partial report condition, like the single letter condition, and the individuals with dyslexia still showed poor performance, suggesting that the difference between the visual span profiles of the two groups cannot be attributed to a deficit in the method of reporting.

The decomposition analyses demonstrated that the underlying contributors of visual span profiles are not different between the dyslexia and control groups. In line with the previous findings (He, Legge & Yu, 2013), crowding effect is the largest determinator of visual span. Mislocation errors also affect the shape of the visual span profiles, although the contribution is substantially smaller compared to crowding. Lastly, the effect of acuity is negligible in both groups to account for visual span size. As supported by the decomposition analyses, we found that visual span sizes are correlated with critical spacing values in letter and symbol conditions of the crowding experiment. This finding provides evidence for the idea that visual span is determined in part by crowding.

Visual span sizes and reading speeds obtained by the Flashcard test are shown to have moderate positive correlation. This finding is in line with the previous literature

(Yu et al., 2007). It also provides motivation for studying visual span in dyslexia due to its close connection with reading speed.

Future Directions

The current study presents interesting preliminary findings demonstrating that people with dyslexia have narrower visual span profiles and require larger inter-stimulus spacings for object recognition in crowding settings compared to typical readers.

It is important to note that, in crowding settings, the difference between people with dyslexia and typical readers cannot be interpreted as all-or-none, as the current results showed no difference in the performance of both groups in the grating condition while dyslexia group showed poorer performance in letter and symbol conditions. Future studies should consider the impact of the nature of the visual stimuli, specifically, whether the stimuli require high attentional resources for feature integration.

In addition, this study is the first to show that people with dyslexia have narrower visual span profiles. However, we observed considerable individual differences within the dyslexia group, even though the group average in dyslexia was significantly lower compared to the control group. Future studies are needed to investigate the visual span profiles in dyslexia with a larger participant pool. The visual span measurement may perhaps serve as a tool to identify some individual differences within the dyslexia population.

References

Agaoglu, M. N., & Chung, S. T. (2016). Can (should) theories of crowding be unified? *Journal of Vision*, *16*(15), 10-10.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.).* Arlington, VA: American Psychiatric Publishing.

Anderson, E. J., Dakin, S. C., Schwarzkopf, D. S., Rees, G., & Greenwood, J. A. (2012). The neural correlates of crowding-induced changes in appearance. *Curr Biol, 22*(13), 1199-1206. doi:10.1016/j.cub.2012.04.063

Beckmann, P. J. (1998). *Preneural factors limiting letter identification in central and peripheral vision*. University of Minnesota.

Beckmann, P. J., & Legge, G. E. (1996). Psychophysics of reading—XIV. The page navigation problem in using magnifiers. Vision research, 36(22), 3723-3733.

Bellocchi, S., Massendari, D., Grainger, J., & Ducrot, S. (2019). Effects of inter-character spacing on saccade programming in beginning readers and dyslexics. *Child Neuropsychology*, *25*(4), 482-506.

Bernard, J. B., Aguilar, C., Castet, E. (2016). A new font, specifically designed for peripheral vision, improves peripheral letter and word recognition, but not eye-mediated reading performance. PLoS One,11:1–25.

Bex, P. J., & Dakin, S. C. (2005). Spatial interference among moving targets. Vision research, 45(11), 1385-1398

Bosse, M. L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition*, *104*(2), 198-230.

Bosse, M. L., & Valdois, S. (2009). Influence of the visual attention span on child reading performance: a cross-sectional study. *Journal of Research in Reading*, *32*(2), 230-253.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*(5241), 177-178. doi:10.1038/226177a0

Bouma, H., & Legein, C. P. (1977). Foveal and parafoveal recognition of letters and words by dyslexics and by average readers. *Neuropsychologia*, *15*(1), 69-80.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis, 10*(4), 433-436.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effect models: a tutorial. Journal of cognition, 1(1), 9, 1-20.

Calabrese, A., Bernard, J. B., Faure, G., Hoffart, L., & Castet, E. (2014). Eye movements and reading speed in macular disease: the shrinking perceptual span hypothesis requires and is supported by a mediation analysis. Investigative ophthalmology & visual science, 55(6), 3638-3645.

Calabrese, A., Bernard, J. B., Faure, G., Hoffart, L., & Castet, E. (2016). Clustering of eye fixations: A new oculomotor determinant of reading speed in maculopathy. Investigative ophthalmology & visual science, 57(7), 3192-3202.

Calabrese, A., To, L., He, Y., Berkholtz, E., Rafian, P., Legge, G. E. (2018). Comparing performance on the MNREAD iPad app with the MNREAD acuity chart. JVis, 18(1):8.

Carlson, T. A., & He, S. (2000). Visible binocular beats from invisible monocular stimuli during binocular rivalry. *Curr Biol, 10*(17), 1055-1058. doi:10.1016/s0960-9822(00)00672-2

Carroll, J. M., Solity, J., & Shapiro, L. R. (2016). Predicting dyslexia using prereading skills: the role of sensorimotor and cognitive abilities. *Journal of Child Psychology and Psychiatry*, *57*(6), 750-758.

Carver, R. P. (1976). Word length, prose difficulty, and reading rate. Journal of Reading Behavior, 8(2), 193-203.

Castet, E., Descamps, M., Denis-Noël, A., & Cole, P. (2017). Letter and symbol identification: No evidence for letter-specific crowding mechanisms. *Journal of Vision*, *17*(11), 2-2.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, *47*(2), 149-180.

Castles, A., Coltheart, M., Larsen, L., Jones, P., Saunders, S., & McArthur, G. (2009). The Castles and Coltheart Test 2 (CC2).

Cestnick, L. (2001). Cross-modality temporal processing deficits in developmental phonological dyslexics. *Brain and Cognition*, *46*(3), 319-325.

Chan, T., Friedman, D. S., Bradley, C., & Massof, R. (2018). Estimates of incidence and prevalence of visual impairment, low vision, and blindness in the United States. JAMA ophthalmology, 136(1), 12-19.

Chen, C., Schneps, M. H., Masyn, K. E., & Thomson, J. M. (2016). The effects of visual attention span and phonological decoding in reading comprehension in dyslexia: a path analysis. *Dyslexia*, *22*(4), 322-344.

Cheong, A. M. Y., Legge, G. E., Lawrence, M., Cheung, S. H., Ruff, M. (2008). Relationship between visual span and reading performance in age-related macular degeneration. Vis Res, 48:577–588.

Cheung, S. H., Kallie, C. S., Legge, G. E., & Cheong, A. M. (2008). Nonlinear mixed-effects modeling of MNREAD data. Investigative ophthalmology & visual science, 49(2), 828-835.

Chicherov, V., Plomp, G., & Herzog, M. H. (2014). Neural correlates of visual crowding. *Neuroimage, 93 Pt 1*, 23-31. doi:10.1016/j.neuroimage.2014.02.021

Chung, S. T. (2007). Learning to identify crowded letters: does it improve reading speed? *Vision Res, 47*(25), 3150-3159. doi:10.1016/j.visres.2007.08.017

Cooper, B., Sun, H., & Lee, B. B. (2012). Psychophysical and physiological responses to gratings with luminance and chromatic components of different spatial frequencies. *JOSA A*, *29*(2), A314-A323.

Crossland, M. D., & Rubin, G. S. (2006). Eye movements and reading in macular disease: Further support for the shrinking perceptual span hypothesis. Vision Research, 46(4), 590-597.

Dacey, D. M. (1994). Physiology, morphology and spatial densities of identified ganglion cell types in primate retina. In *Ciba Foundation symposium* (Vol. 184, p. 12).

Dacey, D. M. (2000). Parallel pathways for spectral coding in primate retina. *Annual review of neuroscience*, *23*(1), 743-775.

Dacey, D. M., & Petersen, M. R. (1992). Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. Proceedings of the National Academy of sciences, 89(20), 9666-9670.

Derrington, A. M., & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *J Physiol, 357*, 219-240. doi:10.1113/jphysiol.1984.sp015498

Doron, A., Manassi, M., Herzog, M. H., & Ahissar, M. (2015). Intact crowding and temporal masking in dyslexia. *Journal of Vision*, *15*(14), 13-13.
Eden, G. F., VanMeter, J. W., Rumsey, J. M., Maisog, J. M., Woods, R. P., & Zeffiro, T. A. (1996). Abnormal processing of visual motion in dyslexia revealed by functional brain imaging. *Nature*, *382*(6586), 66-69.
Ehlers, H. (1936). V: The movements of the eyes during reading. *Acta Ophthalmologica, 14*(1-2), 56-63. doi:10.1111/j.1755-3768.1936.tb07306.
Fine, E. M., & Peli, E. (1996). The role of context in reading with central field loss. Optometry and vision science: official publication of the American Academy of Optometry, 73(8), 533-539.
Fischer, J., & Whitney, D. (2011). Object-level visual information gets through the bottleneck of crowding. Journal of Neurophysiology, 106(3), 1389-1398.
Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K., & Facoetti, A. (2012). A causal link between visual spatial attention and reading acquisition. *Current Biology*, *22*(9), 814-819.
Francks, C., MacPhie, I. L., & Monaco, A. P. (2002). The genetic basis of dyslexia. *The Lancet Neurology*, *1*(8), 483-490.
Frey, A., & Bosse, M. L. (2018). Perceptual span, visual span, and visual attention span: Three potential ways to quantify limits on visual processing during reading. *Visual Cognition*, *26*(6), 412-429.
Galaburda, A. M. (1999). Developmental dyslexia: A multilevel syndrome. *Dyslexia*, *5*(4), 183-191.
Gallagher, A., Frith, U., & Snowling, M. J. (2000). Precursors of literacy delay among children at genetic risk of dyslexia. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *41*(2), 203-213.
Germanò, E., Gagliano, A., & Curatolo, P. (2010). Comorbidity of ADHD and dyslexia. *Developmental neuropsychology*, *35*(5), 475-493
Green, P., & MacLeod C. J. (2016) SIMR: an R package for power analysis of generalized linear mixed models by simulation. Methods in Ecology and Evolution, 7, 493-498.
Grainger, J., Tydgat, I., & Issele, J. (2010). Crowding affects letters and symbols differently. *J Exp Psychol Hum Percept Perform, 36*(3), 673-688. doi:10.1037/a0016888
Granquist, C., Wu, Y. H., Gage, R., Crossland, M. D., Legge, G. E. (2018). How people with Low vision achieve magnification in digital reading. Optom Vis Sci, 95(9): 711-719
Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2010). Crowding changes appearance. *Current Biology*, *20*(6), 496-501.
Greenwood, J. A., & Parsons, M. J. (2020). Dissociable effects of visual crowding on the perception of color and motion. Proceedings of the National Academy of Sciences, 117(14), 8196-8202.
Hakvoort, B., van den Boer, M., Leenaars, T., Bos, P., & Tijms, J. (2017). Improvements in reading accuracy as a result of increased interletter spacing are not specific to children with dyslexia. *Journal of experimental child psychology*, *164*, 101-116.
He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383*(6598), 334-337.

He, Y., Legge, G. E., & Yu, D. (2013). Sensory and cognitive influences on the training-related improvement of reading speed in peripheral vision. *Journal of vision*, *13*(7), 14-14.

Herzog, M. H., & Manassi, M. (2015). Uncorking the bottleneck of crowding: A fresh look at object recognition. Current Opinion in Behavioral Sciences, 1, 86-93.

Jednoróg, K., Gawron, N., Marchewka, A., Heim, S., & Grabowska, A. (2014). Cognitive subtypes of dyslexia are characterized by distinct patterns of grey matter volume. *Brain Structure and Function*, *219*(5), 1697-1707.

Joanisse, M. F., Manis, F. R., Keating, P., & Seidenberg, M. S. (2000). Language deficits in dyslexic children: Speech perception, phonology, and morphology. *Journal of experimental child psychology*, *77*(1), 30-60.

Joo, S. J., White, A. L., Strodtman, D. J., & Yeatman, J. D. (2018). Optimizing text for an individual's visual system: The contribution of visual crowding to reading difficulties. *Cortex*, *103*, 291-301.

Kennedy, G. J., & Whitaker, D. (2010). The chromatic selectivity of visual crowding. *Journal of Vision*, *10*(6), 15-15.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spat Vis, 8*(2), 255-279. doi:10.1163/156856894x00350

Kwon, M., & Liu, R. (2019). Linkage between retinal ganglion cell density and the nonuniform spatial integration across the visual field. Proceedings of the National Academy of Sciences, 116(9), 3827-3836.

Lachmann, T., Berti, S., Kujala, T., & Schröger, E. (2005). Diagnostic subgroups of developmental dyslexia have different deficits in neural processing of tones and phonemes. *International Journal of Psychophysiology*, *56*(2), 105-120.

Lee, B. B., Pokorny, J., Smith, V. C., Martin, P. R., & Valbergt, A. (1990). Luminance and chromatic modulation sensitivity of macaque ganglion cells and human observers. *JOSA A*, *7*(12), 2223-2236.

Legge, G. E. (2007). Psychophysics of reading in normal and low vision. Mahwah, NJ and London: Lawrence Erlbaum Associates.

Legge, G. E. (2016). Reading digital with low vision. Visible language, 50(2), 102

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), 8-8.

Legge, G. E., Mansfield, J. S., & Chung, S. T. (2001). Psychophysics of reading: XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision research*, *41*(6), 725-743.

Legge, G. E., Pelli, D. G., Rubin, G. S., & Schleske, M. M. (1985). Psychophysics of reading—I. Normal vision. Vision research, 25(2), 239-252.

Legge, G. E., Ross, J. A., Luebker, A., & Lamay, J. M. (1989). Psychophysics of reading. VIII. The Minnesota low-vision reading test. Optom Vis Sci, 66(12), 843-53.

Leonova, A., Pokorny, J., & Smith, V. C. (2003). Spatial frequency processing in inferred PC-and MC-pathways. Vision Research, 43(20), 2133-2139.

Letters between friends. (2017). University of Louisville Magazine, 36:28–31.

Lev, M., Yehezkel, O., & Polat, U. (2014). Uncovering foveal crowding?. Scientific reports, 4, 4067.

Levi, D. M. (2008). Crowding--an essential bottleneck for object recognition: a mini-review. *Vision Res, 48*(5), 635-654. doi:10.1016/j.visres.2007.12.009
Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Res, 25*(7), 963-977. doi:10.1016/0042-6989(85)90207-x
Levi, D. M., Song, S., & Pelli, D. G. (2007). Amblyopic reading is crowded. *J Vis, 7*(2), 21 21-17. doi:10.1167/7.2.21
Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *J Neurosci, 7*(11), 3416-3468.
Lobier, M., Zoubrinetzky, R., & Valdois, S. (2012). The visual attention span deficit in dyslexia is visual and not verbal. cortex, 48(6), 768-773.
Lorusso, M. L., Cantiani, C., & Molteni, M. (2014). Age, dyslexia subtype and comorbidity modulate rapid auditory processing in developmental dyslexia. *Frontiers in human neuroscience*, *8*, 313.
Louie, E. G., Bressler, D. W., & Whitney, D. (2007). Holistic crowding: Selective interference between configural representations of faces in crowded scenes. Journal of Vision, 7(2), 24-24.
Manassi, M., Sayim, B., & Herzog, M. H. (2012). Grouping, pooling, and when bigger is better in visual crowding. *J Vis, 12*(10), 13. doi:10.1167/12.10.13
Manassi, M., & Whitney, D. (2018). Multi-level crowding and the paradox of object recognition in clutter. Current Biology, 28(3), R127-R133.
Mansfield, J. S., & Legge, G.E. (2007). The MNREAD acuity chart. In Psychophysics of Reading in Normal and Low Vision. Lawrence Erlbaum Assoc, Mahwah, NJ/London, Ch.5, pp. 167–91.
Mansfield, J. S., Legge, G. E., & Bane, M. C. (1996). Psychophysics of reading. XV: Font effects in normal and low vision. Investigative Ophthalmology & Visual Science, 37(8), 1492-1501.
Martelli, M., Di Filippo, G., Spinelli, D., & Zoccolotti, P. (2009). Crowding, reading, and developmental dyslexia. *Journal of vision*, *9*(4), 14-14.
McAnany, J. J., & Levine, M. W. (2007). Magnocellular and parvocellular visual pathway contributions to visual field anisotropies. Vision Research, 47(17), 2327-2336.
Merigan, W. H., Byrne, C. E., & Maunsell, J. H. (1991). Does primate motion perception depend on the magnocellular pathway? *J Neurosci, 11*(11), 3422-3429.
McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. Perception & Psychophysics, 17(6), 578-586.
Miellet, S., O'Donnell, P. J., & Sereno, S. C. (2009). Parafoveal magnifi-cation: Visual acuity does not modulate the perceptual span in reading. Psychological Science, 20, 721–728.
Miles, T. R. (2004). Some problems in determining the prevalence of dyslexia.
Millin, R., Arman, A. C., Chung, S. T., & Tjan, B. S. (2014). Visual crowding in V1. *Cereb Cortex, 24*(12), 3107-3115. doi:10.1093/cercor/bht159.
Moores, E., Cassim, R., & Talcott, J. B. (2011). Adults with dyslexia exhibit large effects of crowding, increased dependence on cues, and detrimental effects of distractors in visual search tasks. *Neuropsychologia*, *49*(14), 3881-3890.

Morrison, R. E., & Rayner, K. (1981). Saccade size in reading depends upon character spaces and not visual angle. Perception & psychophysics.

Nassi, J. J., & Callaway, E. M. (2009). Parallel processing strategies of the primate visual system. *Nat Rev Neurosci, 10*(5), 360-372. doi:10.1038/nrn2619.

O'Regan, J. K. (1983). Elementary perceptual and eye movement control processes in reading. In K. Rayner (Ed.), Eye movements in reading: Perceptual and language processes (pp. 121–139). New York, NY: Academic Press. http://dx.doi.org/10.1016/B978-0-12-583680-7 .50012-6

Parker, A. J., Slattery, T., J., & Kirkby, J. A. (2019). Return-sweep saccades during reading in adults and children. Vision research, 155, 35-43.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis, 10*(4), 437-442.

Pelli, D. G. (2008). Crowding: a cortical constraint on object recognition. *Curr Opin Neurobiol, 18*(4), 445-451. doi:10.1016/j.conb.2008.09.008

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: distinguishing feature integration from detection. *J Vis, 4*(12), 1136-1169. doi:10.1167/4.12.12

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *J Vis, 7*(2), 20 21-36. doi:10.1167/7.2.20.

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.

Perrachione, T. K., Del Tufo, S. N., Winter, R., Murtagh, J., Cyr, A., Chang, P., … & Gabrieli, J. D. (2016). Dysfunction of rapid neural adaptation in dyslexia. *Neuron*, *92*(6), 1383-1397.

Pokorny, J., & Smith, V. C. (1997). Psychophysical signatures associated with magnocellular and parvocellular pathway contrast gain. JOSA A, 14(9), 2477-2486.

Põder, E. (2007). Effect of colour pop-out on the recognition of letters in crowding conditions. Psychological Research, 71(6), 641-645.

Ramus, F., & Ahissar, M. (2012). Developmental dyslexia: The difficulties of interpreting poor performance, and the importance of normal performance. *Cognitive neuropsychology*, *29*(1-2), 104-122.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. Journal of experimental child psychology, 41(2), 211-236.

Rayner, K. (1998). Eye movement in reading and information processing: 20 years of research. Psychological Bulletin, 124(3), 372-422.

Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. Science, 206(4417), 468-469.

Rayner, K., Slattery, T. J., & Bélanger, N. N. (2010). Eye movements, the perceptual span, and reading speed. Psychonomic bulletin & review, 17(6), 834-839.

Rodieck, R. W., Binmoeller, K. F., & Dineen, J. (1985). Parasol and midget ganglion cells of the human retina. *Journal of Comparative Neurology*, *233*(1), 115-132.

Sayim, B., Westheimer, G., & Herzog, M. H. (2008). Contrast polarity, chromaticity, and stereoscopic depth modulate contextual interactions in vernier acuity. Journal of Vision, 8(8), 12-12.

Schiller, P. H., Logothetis, N. K., & Charles, E. R. (1990). Functions of the colour-opponent and broad-band channels of the visual system. *Nature, 343*(6253), 68-70. doi:10.1038/343068a0.

Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine*, *326*(3), 145-150.

Shovman, M. M., & Ahissar, M. (2006). Isolating the impact of visual perception on dyslexics' reading ability. *Vision Research*, *46*(20), 3514-3525.

Stein, J. (2001). The magnocellular theory of developmental dyslexia. *Dyslexia, 7*(1), 12-36. doi:10.1002/dys.186

Stein, J. (2014). Dyslexia: the role of vision and visual attention. *Current developmental disorders reports*, *1*(4), 267-280.

Tapia, E., & Breitmeyer, B. G. (2011). Visual consciousness revisited: magnocellular and parvocellular contributions to conscious and nonconscious vision. *Psychol Sci, 22*(7), 934-942. doi:10.1177/0956797611413471

Text Readability Consensus Calculator (n.d.). Readability Formulas. Retrieved March 13, 2020, from https://readabilityformulas.com/freetests/six-readability-formulas.php.

Tinker M. A. (1963) Legibility of Print. Iowa State University Press, Ames.

Tinker, M. A., & Paterson, D. G. (1931). Studies of typographical factors influencing speed of reading. V. Simultaneous variation of type size and line length. Journal of Applied Psychology, 15(1), 72.

The Brothers Grimm (2008) Grimms' Fairy Tales. Retrieved from http://www.gutenberg.org/ebooks/2591

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Res, 32*(7), 1349-1357. doi:10.1016/0042-6989(92)90227-a

Tripathy, S. P., Cavanagh, P., & Bedell, H. E. (2014). Large crowding zones in peripheral vision for briefly presented stimuli. Journal of Vision, 14(6), 11-11.

Tydgat, I., & Grainger, J. (2009). Serial position effects in the identification of letters, digits, and symbols. *J Exp Psychol Hum Percept Perform, 35*(2), 480-498. doi:10.1037/a0013027.

van Ermingen-Marbach, M., Grande, M., Pape-Neumann, J., Sass, K., & Heim, S. (2013). Distinct neural signatures of cognitive subtypes of dyslexia with and without phonological deficits. *NeuroImage: Clinical*, *2*, 477-490.

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.

Visser, T. A., Boden, C., & Giaschi, D. E. (2004). Children with dyslexia: evidence for visual attention deficits in perception of rapid sequences of objects. *Vision Research*, *44*(21), 2521-2535.

Warmington, Meesha (2016): York Adult Assessment Battery Revised Form and Test.pdf. figshare. Journal contribution. https://doi.org/10.6084/m9.figshare.2062560.v1

Wen, W., Zhang, P., Liu, T., Zhang, T., Gao, J., Sun, X., & He, S. (2015). A Novel Motion-on-Color Paradigm for Isolating Magnocellular Pathway Function in Preperimetric Glaucoma. *Invest Ophthalmol Vis Sci, 56*(8), 4439-4446. doi:10.1167/iovs.15-16394.

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.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. Perception & Psychophysics, 63(8), 1293–1313.

Williams, M. J., Stuart, G. W., Castles, A., & McAnally, K. I. (2003). Contrast sensitivity in subgroups of developmental dyslexia. *Vision research*, *43*(4), 467-477.

Wu, Y. H., Granquist, C., Gage, R., Crossland, M. D., & Legge, G. E. (2020). Online Survey of Digital Reading by Adults with Low Vision. Optometry and Vision Science, 97(4), 249-256.

Xiong, Y. Z., Calabrese, A., Cheong, A. M. Y., Legge, G. E. (2018). Reading acuity as a predictor of low-vision reading performance. Invest Ophthalmol Vis Sci, 59, 4798–4803.

Xiong, Y. Z., Lorsung, E. A., Mansfield, J. S., Bigelow, C., Legge, G. E. (2018). Fonts designed for macular degeneration: impact on reading. Invest Ophthalmol Vis Sci, 8(59): 4182–4189.

Yu, D., Cheung, S. H., Legge, G. E., & Chung, S. T. (2007). Effect of letter spacing on visual span and reading speed. *Journal of vision*, *7*(2), 2-2.

Ziegler, J. C., Pech‐Georgel, C., Dufau, S., & Grainger, J. (2010). Rapid processing of letters, digits and symbols: what purely visual‐attentional deficit in developmental dyslexia?. *Developmental science*, *13*(4), F8-F14.

Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M., & Ziegler, J. C. (2012). Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences*, *109*(28), 11455-11459

Appendix 1: Individual reading curves of normally sighted subjects in Chapter 2

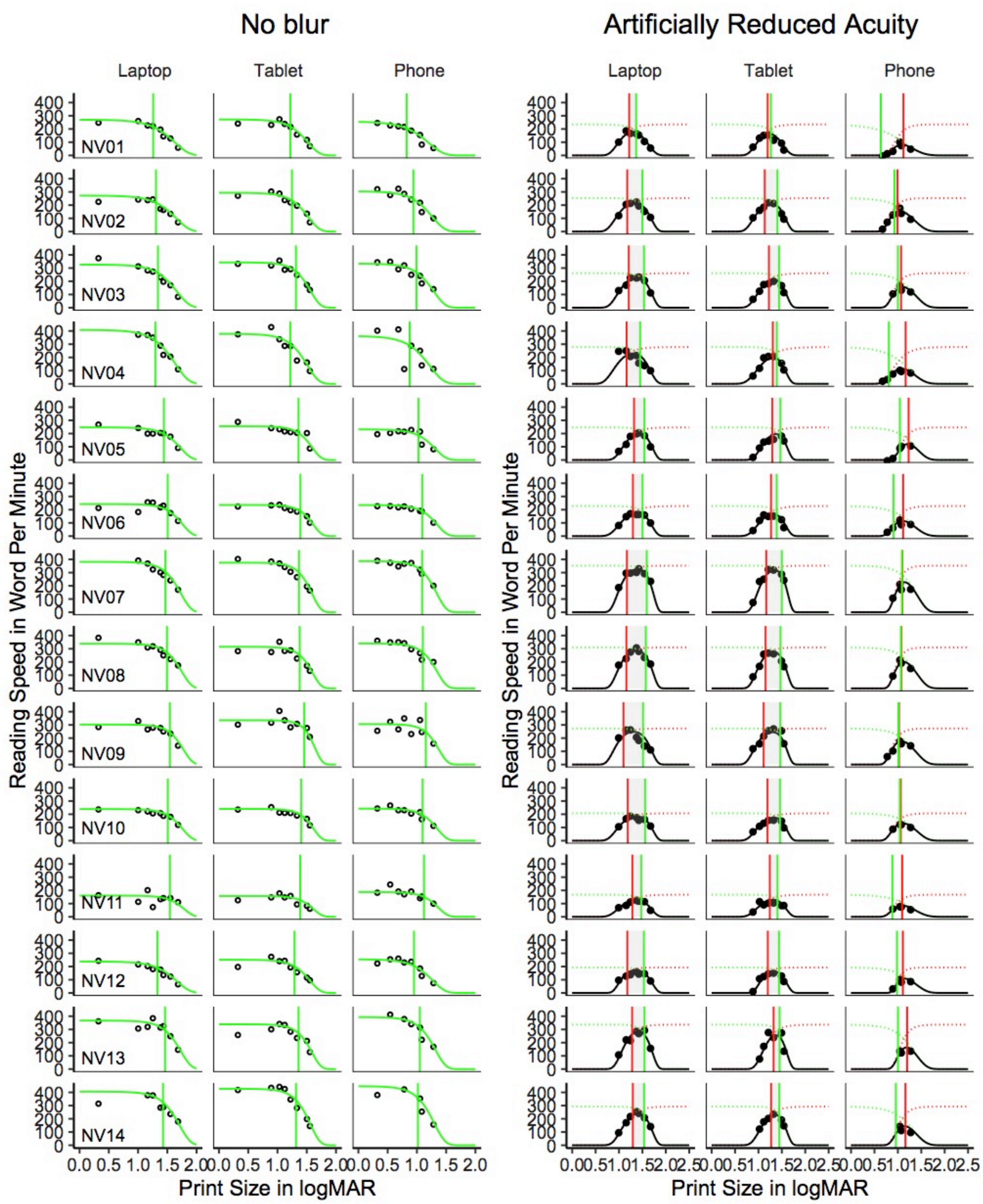


Figure A1- 1: Individual reading curves of normally sighted subjects reading with Times New Roman font. Reading speed (words per minute is plotted as a function of print size (logMAR units) for 14 subjects with normal vision (NV01 – NV14). Data are shown for display formats simulating a cellphone, tablet and laptop. Left: no blur. Right: artificially reduced acuity using diffusive blur.

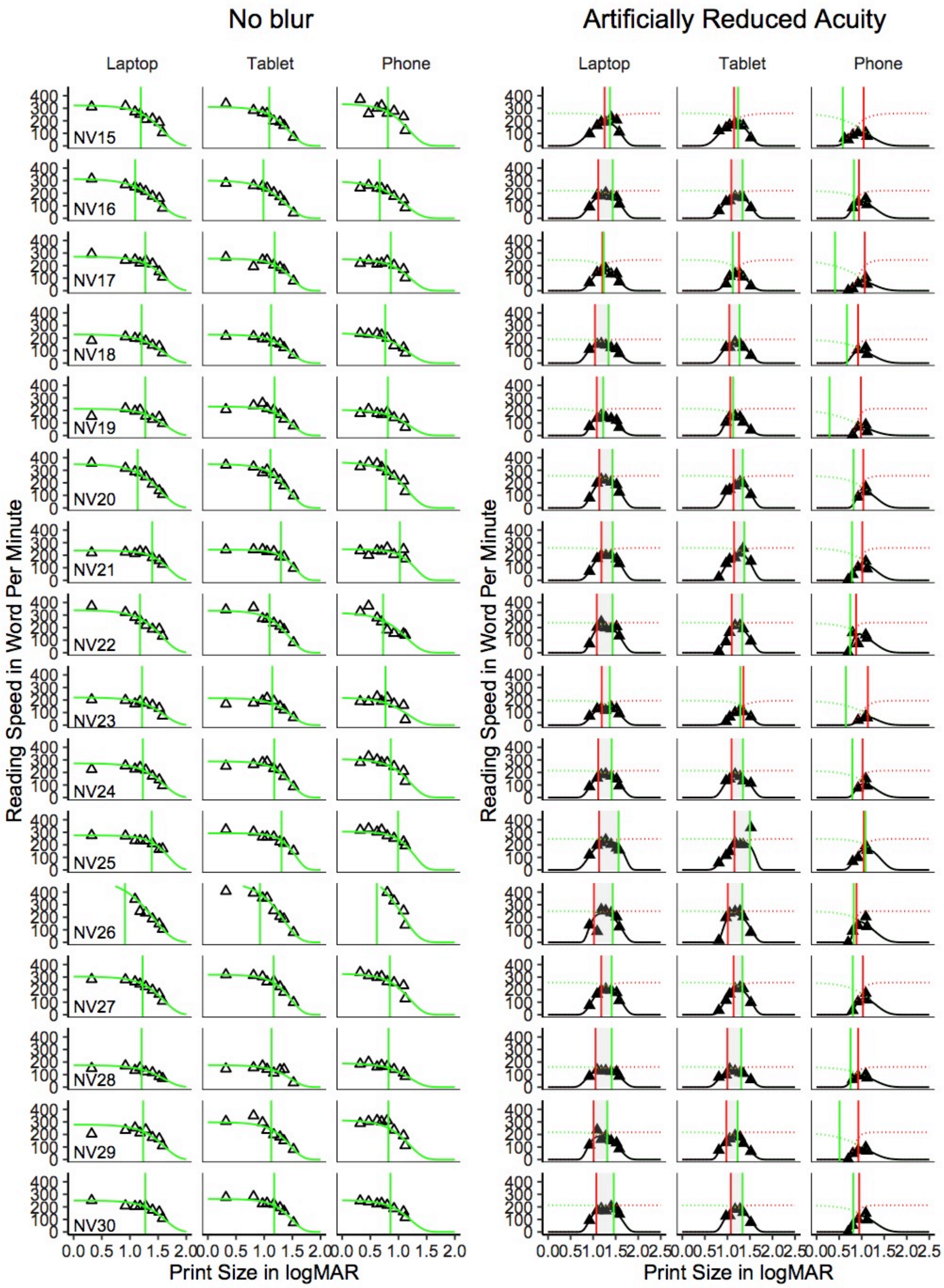


Figure A1- 2: Individual reading curves of normally sighted subjects reading with Courier font. Reading speed (words per minute is plotted as a function of print size (logMAR units) for 16 subjects with normal vision (NV15 – NV30). Data are shown for display formats simulating a cellphone, tablet and laptop. Left: no blur. Right: artificially reduced acuity using diffusive blur.

Appendix 2. Estimating the Number of Characters per Line in Chapter 2

Estimating the Number of Characters per Line from Angular Print Size

For any display width W (in cm), the number of characters N on each line can be calculated based on the angular print size P and vice versa. For continuous text, the character size can be defined as the height of a lower case "x" (x-height), and expressed in snellen or logMAR units. Here we adopted logMAR units. An x-height of 0 logMAR corresponds to 0.083 degree, and each 1 logMAR increase is associated with 10 times increase in angular size. Therefore, for any x-height P, the corresponding size in radians can be obtained by $0.083*10^P/180*\pi$, which can be simplified as $0.00145*10^P$. The corresponding average character width in radians:

$$Average\ character\ width\ in\ radians = 0.00145\ 10^{P}\ K; \qquad \text{(Eq. S1)}$$

where the constant K is the ratio between average character width and x-height, which depends on the font, e.g., K is 0.96 for Times and 1.33 for Courier.

For a display width W in cm, its angular width in radians depends on the viewing distance V (in cm):

$$Angular\ line\ width\ in\ radians \approx \frac{W}{V}; \qquad \text{(Eq. S2)}$$

Therefore, the character number per line can be obtained by:

$$N = \frac{Angular\ line\ width\ in\ radians}{Average\ character\ width\ in\ radians} \approx \frac{W}{0.00145\ 10^{P}\ K\ V}; \qquad \text{(Eq. S3)}$$

To verify this calculation process, we compared the estimated character per line with the actual character per line for the character sizes we tested on each display format. As shown in Figure S3, there is good agreement between the estimated and actual character counts.

Similarly, when N is known, the corresponding angular print size P can be obtained by:

$$P = log10\left(\frac{W}{0.00145\ K\ V\ N}\right); \qquad \text{(Eq. S5)}$$

And when both N and P are known, the smallest display width W can be obtained by:

$$W = 0.00145\ 10^{P}\ K\ V\ N. \qquad \text{(Eq. S6)}.$$

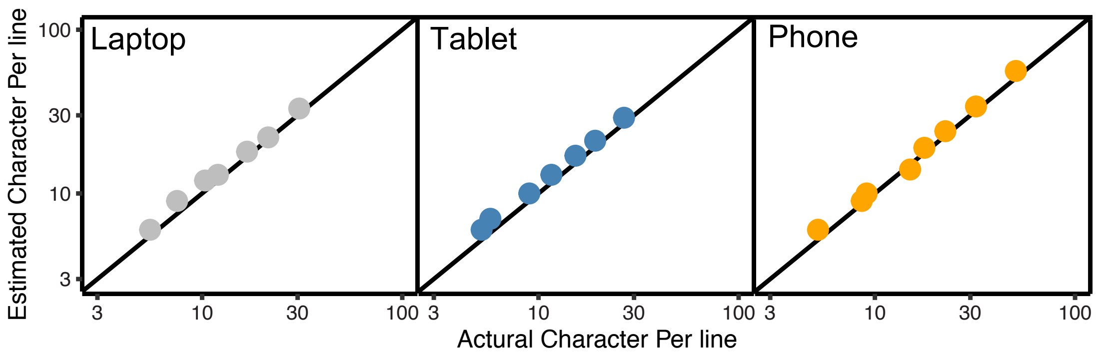


Figure A2- 1: Comparison between the estimated and actual number of characters per line, for each of the three displays. The estimated number of characters per line was calculated using equation Eq.S3 based on the logMAR print size, and the actual number of characters per line was obtained from the character counting function of Microsoft Word 2016.

Appendix 3: Relationship Between Critical Print Size and Visual Acuity in Chapter 2

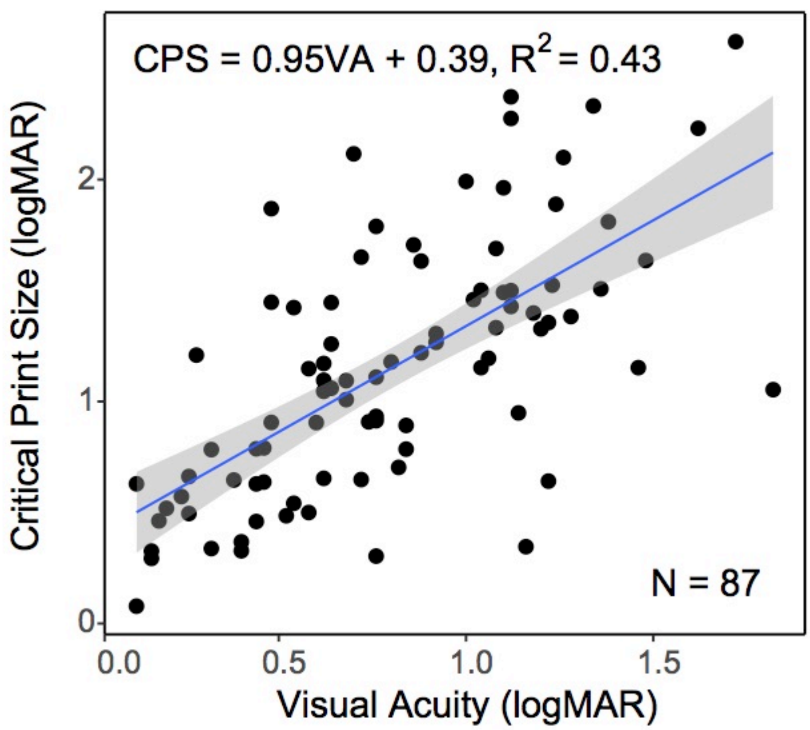


Figure A3- 1: Critical print size (CPS) as a function of visual acuity (VA). The CPS was measured with the MNREAD acuity charts, and VA was measured with the ETDRS letter acuity charts. Data from 87 low vision participants are shown, including 77 participants from published studies in our lab (29-31), and 10 participants who participated in the current study. The regression line and linear regression function were provided in the figure.

Appendix 4: Font Properties in Chapter 2

Font	K (Width/ X-height)	MDW, 15cm, 0.5 logMAR	MDW, 15cm, 1.0 logMAR	MDW, 15cm, 1.5 logMAR	MDW, 40cm, 0.5 logMAR	MDW, 40cm, 1.0 logMAR	MDW, 40cm, 1.5 logMAR
Arial	0.94	1.4	4.1	12.3	3.7	10.9	32.8
Arial Bold	1.04	1.4	4.2	12.6	3.7	11.2	33.6
Arial Italic	0.94	1.4	4.1	12.3	3.7	10.9	32.8
Calibri	0.98	1.4	4.1	12.4	3.7	11.1	33.1
Calibri Bold	1.00	1.4	4.2	12.5	3.7	11.1	33.3
Calibri Italic	0.97	1.4	4.1	12.4	3.7	11.0	33.0
Courier	1.33	1.4	4.3	12.9	3.8	11.5	34.5
Courier Bold	1.33	1.4	4.3	12.9	3.8	11.5	34.5
Courier Italic	1.33	1.4	4.3	12.9	3.8	11.5	34.5
Eido	0.98	1.4	4.1	12.4	3.7	11.0	33.1
Eido Mono	1.21	1.4	4.3	12.9	3.8	11.5	34.3
Helvetica	0.96	1.4	4.1	12.3	3.7	11.0	32.9
Helvetica Bold	1.01	1.4	4.2	12.5	3.7	11.1	33.4
Helvetica Italic	0.96	1.4	4.1	12.3	3.7	11.0	32.9
Maxular	1.54	1.4	4.3	12.8	3.8	11.3	34.0
Times	0.96	1.4	4.1	12.3	3.7	11.0	32.9
Times Bold	1.07	1.4	4.2	12.7	3.8	11.3	33.8
Times Italic	1.04	1.4	4.2	12.6	3.7	11.2	33.6

Table A4- 1: Font Properties and the Corresponding Minimum Display Width (MDW) for Hypothetical Low Vision Subjects with Mild, Moderate, and Severe Vision Loss. For Columns 3-8: Minimum Display Width (cm), viewing distance (cm), visual acuity (logMAR).

Appendix 5. Character Counts and Print Sizes in Chapter 2

Display Format	Font	Mean Character Count Per Page (Range)	Print Size, logMAR (Range) *	Print Size, pts (Range)
Phone (2.3" x 4.1")	Times New Roman	557 - 5.2	0.34 - 1.3	12 - 108
	Courier	398 – 5.2	0.32 – 1.12	12 - 76
Tablet (5.9" x 7.9")	Times New Roman	3090 – 5.2	0.34 - 1.57	12 - 204
	Courier	2165 – 3.8	0.32 – 1.52	12 - 190
Laptop (6.6" x 11.4")	Times New Roman	5265 – 5.5	0.34 - 1.69	12 - 276
	Courier	3511 – 5.2	0.32 – 1.58	12 - 216

Table A5- 1: Character Counts and Print Sizes for Three Display Formats and two Fonts. * The print sizes in logMAR units shown here are for a viewing distance of 60 cm. At the 40 cm viewing distance used by the low vision subjects, the corresponding logMAR print sizes were approximately 0.18 log units larger.

Appendix 6: The individual scores in Assessments and Reading Speed Test in Chapter 4

		Assessments											
		Flashcard Test	Test of Auditory Processing (in percentile rank)			Woodcock Johnson IV Test of Achievement (in percentile rank)					Castle & Coltheart Test 2 (in z-scores)		
group	age	Reading Speed (word per min)	Word discrimination	Phoneme segmentation	Phonological blending	Letter-Word recognition	Word attack	spelling	Word reading fluency	Sentence reading fluency	CC2_ regular	CC2_ irregular	CC2_ nonwords
dyslexia	60	371	75	16	16	79	35	66	38	45	1.1	2.58	-0.8
dyslexia	27	343	50	37	91	50	35	33	19	3	1.1	2.58	-1.59
dyslexia	54	264	16	1	2	8	3	1	6	5	-1.92	-0.87	-2.13
dyslexia	29	252	75	16	16	14	25	17	14	12	-1.43	-0.15	-0.87
dyslexia	34	375	50	2	25	NA	NA	NA	NA	NA	0.62	1.43	-1.66
dyslexia	32	273	75	25	25	50	42	68	66	66	1.1	2.07	-0.74
dyslexia	36	365	75	37	9	30	4	NA	9	9	-0.62	1.43	-1.62
dyslexia	28	406	75	9	9	45	12	53	70	66	0.62	1.43	-0.94
dyslexia	18	296	75	25	75	18	18	18	19	45	0.62	0.74	-1.2
dyslexia	19	342	25	25	16	18	19	19	18	19	-0.24	1.43	-0.65
dyslexia	28	297	75	25	16	55	38	38	19	63	1.59	1.43	-0.65
control	20	654	75	50	75	50	30	50	88	91	0.62	2.58	-0.29
control	32	428	50	50	75	63	60	33	50	53	1.1	2.58	-0.67
control	51	653	75	50	25	63	45	50	83	91	NA	NA	NA
control	29	453	75	75	91	72	76	72	68	70	1.1	2.58	-0.65
control	22	562	75	25	75	75	72	75	78	78	1.59	2.58	-0.62
control	29	624	75	75	95	94	76	91	78	75	1.59	2.58	-0.29
control	19	595	75	50	27	76	75	72	75	83	1.59	2.58	-0.29
control	20	495	75	37	16	55	63	24	53	50	-0.62	2.58	-0.94
control	45	572	75	75	50	84	84	91	84	86	1.1	2.58	-0.62
control	19	350	75	2	16	17	18	33	70	75	1.1	2.07	-0.77

Table A6- 1: Individual scores in all of the standardized assessments and Flashcard reading speed test.

Appendix 7: Individual Critical Spacing Values in Chapter 4

SubjectID	Group	Critical Spacing (visual degrees)		
		letter	symbol	grating
10	control	9.1	10.6	12.5
12	control	3.83	3.55	5.8
13	control	8.3	8.5	11.7
14	control	5.9	6.3	8.3
17	control	3.8	5.27	3.7
19	control	9.2	8.95	9.9
22	control	5.3	4.5	6
23	control	8.6	9	14
8	control	5.7	3.9	8.75
9	control	6.2	5.6	10.6
1	dyslexia	10.3	7.5	11.3
11	dyslexia	9.5	11	10.1
16	dyslexia	6.6	8.5	11
2	dyslexia	7.7	9.5	11
20	dyslexia	9.3	6.45	11.1
21	dyslexia	6.45	10.1	8.3
3	dyslexia	10.7	9.5	14
4	dyslexia	8.18	6.7	7.87
5	dyslexia	5.9	8.6	10
6	dyslexia	9.3	9.5	12.75
7	dyslexia	8.3	7.2	12.1

Table A7- 1: Individual critical spacing values for all the participants in each stimulus category in the crowding experiment.

Appendix 8: Correlation Coefficients for all the measurements in Chapter 4

	Word discrimination	Phoneme segmentation	Phonological blending	CC2_ regular	CC2_ irregular	CC2_ nonwords	Letter-Word recognition	Word attack	Spelling	Word reading fluency	Sentence reading fluency	Letter critical spacing	Symbol critical spacing	Grating critical spacing	Visual span size	Reading speed (wpm)
Word discrimination	1															
Phoneme segmentation	0.28	1														
Phonological blending	0.16	0.65	1													
CC2_regular	0.54	0.45	0.44	1												
CC2_irregular	0.47	0.61	0.45	0.77	1											
CC2_nonwords	0.6	0.54	0.18	0.62	0.64	1										
Letter-Word recognition	0.48	0.72	0.43	0.7	0.79	0.57	1									
Word attack	0.41	0.81	0.48	0.52	0.69	0.55	0.85	1								
Spelling	0.57	0.62	0.38	0.71	0.68	0.63	0.87	0.74	1							
Word reading fluency	0.58	0.53	0.26	0.59	0.7	0.7	0.63	0.58	0.76	1						
Sentence reading fluency	0.68	0.48	0.19	0.69	0.63	0.78	0.6	0.53	0.71	0.9	1					
Letter critical spacing	-0.35	-0.76	-0.61	-0.46	-0.39	-0.48	-0.46	-0.63	-0.45	-0.36	-0.45	1				
Symbol critical spacing	-0.06	-0.66	-0.47	-0.07	-0.29	-0.43	-0.34	-0.5	-0.36	-0.29	-0.23	0.71	1			
Grating critical spacing	-0.38	-0.65	-0.26	-0.22	-0.26	-0.46	-0.49	-0.61	-0.47	-0.26	-0.31	0.81	0.66	1		
Visual span size	0.18	0.67	0.51	0.4	0.59	0.37	0.41	0.6	0.54	0.58	0.37	-0.51	-0.49	-0.3	1	
Reading speed (wpm)	0.37	0.69	0.38	0.48	0.68	0.66	0.69	0.64	0.62	0.81	0.73	-0.44	-0.39	-0.39	0.43	1

Table A8- 1: Correlation coefficients for all the measurements including standardized assessments and psychophysical experiments.