The Influence of Stroke Width on Legibility for Low Vision Adults: Integrating Scientific & Design Knowledge on Typeface Boldness

Karin von Ompteda

A thesis submitted in partial fulfilment of the requirements of the Royal College of Art for the degree of Doctor of Philosophy

April 2022

COPYRIGHT

AUTHOR'S DECLARATION

Some of the contents of Chapters 1-3 have previously been published in a paper for the *Include 2009* conference entitled 'Innovation in Inclusive Typography: A Role for Design Research' (von Ompteda, 2009a). Further, some of the contents of Chapters 3, 5, and 6 have previously been published in an article for *Grafik Magazine* entitled 'Typo Transparencies' (von Ompteda, 2011a). The reviews and analyses presented in this thesis are significantly more extensive, as well as updated to 2021.

This thesis represents partial submission for the degree of Doctor of Philosophy at the Royal College of Art. I confirm that the work presented here is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

During the period of registered study in which this thesis was prepared the author has not been registered for any other academic award or qualification. The material included in this thesis has not been submitted wholly or in part for any academic award or qualification other than that for which it is now submitted.

Karin von Ompteda
April 2022

ABSTRACT

This PhD thesis investigates the influence of typeface stroke width on reading performance for low vision adults. While scientific evidence suggests that an increased stroke width—or bolder typeface—can improve legibility, optimal values are not well understood. In keeping with this, existing accessible design guidelines in the United Kingdom recommend a large range of typeface weights from regular to bold. The goal of this PhD research is to inform print design guidelines with a higher degree of specificity, and thereby increase the proportion of the population able to access text.

This research is based upon an initial inquiry formulated around one main question: What is the optimal typeface stroke width for low vision adults? In order to address this question, an integration of knowledge drawing from vision science and typographic design is undertaken. The majority of research into typeface legibility exists within vision science, while the creation of typefaces and expertise in their use exists within the discipline of design. This PhD responds to the lack of interdisciplinary approaches to typeface legibility research, which has resulted in limited application of scientific research to design practice.

This practice-based communication design PhD addresses the research question through a quantitative analysis of text typefaces. This involves the measurement of typeface proportions and the analysis of this typeface data through information visualisation. Typeface data is initially gathered with the purpose of designing a typeface for experimental testing. It is through this typeface design practice that the methods for the quantitative analysis of typefaces emerge, which then become the focus of the research. This PhD investigation develops a foundation of interdisciplinary—science and design—typographic knowledge, based on typeface data.

This research consolidates scientific knowledge on the influence of boldness on legibility in the context of low vision. Ten scientific legibility studies are analysed. This entails measuring and visualising the stroke width values of typefaces that have been experimentally found to have higher and lower legibility.

Design knowledge is formalised by measuring and visualising the stroke width values of typefaces commonly used in design practice. This is a *design phenomenology* study as defined by Nigel Cross, investigating design knowledge residing in artefacts themselves. By integrating scientific and design knowledge as proposed, interdisciplinary knowledge on typeface legibility for low vision adults is developed. My original contribution to knowledge includes visualising how the stroke widths of typefaces experimentally found to improve legibility relate to the stroke widths of typefaces commonly employed in design practice.

This thesis concludes that typefaces with stroke width values ranging from 22-33% (percent of x-height) improve legibility in the context of low vision. The analysis further indicates that sans serif regular typefaces range from 13.5-19.8% stroke width and are not optimal for low vision reading. The analysis also indicates that sans serif bold typefaces range from 18.9-40.0% stroke width, and that many, but not all, may improve reading performance for adults with low vision. This research is intended to be useful for legibility researchers and the development of evidence-based accessible design guidelines.

ACKNOWLEDGMENTS

My deepest gratitude to my supervisor Professor Teal Triggs for her unwavering support of my work, and for helping me to become the researcher I am today. I am grateful to my first supervisors Jeff Willis and Tony Cobb, for their support of my research project and career. Many thanks to Dr Aries Arditi for his external supervision early in the research process. My sincere appreciation for Professor Paul Luna whose feedback made an enormous impact on the final thesis. Thank you to the many staff at the Royal College of Art who have supported my PhD research, including Dr Catherine Dormor, Professor Jeremy Aynsley, Professor Dan Fern, and Tom Simmons. My heartfelt gratitude for the late Al Rees, my dear comrade. Special thanks to Adrian Shaughnessy and Professor Neville Brody for their contributions to my professional development. I thank my OCAD University colleagues who have supported me while I balanced PhD completion alongside my professor role.

Without funding, this PhD would not have been possible. I extend my profound appreciation to the Commonwealth Scholarship Commission in the United Kingdom, the Social Sciences and Humanities Research Council of Canada (SSHRC), the Royal College of Art, and the Canadian Centennial Scholarship Fund.

A wonderful community of friends and family have supported me through this long journey. Thank you to my family, Ingrid and Fred von Ompteda, Robert and Hildegard Misch, and Karina and Len Logozar. Thank you to my dear friends, Povilas Utovka, Katie Gaudion, and Natasha Myers. Many thanks to Peter Crnokrak for his support in the early years of the research. My sincere gratitude to Saadia Kardar for endless practical assistance and ceaseless encouragement. I am eternally grateful to my partner John Obercian for his infinite love and support.

I dedicate my PhD to Professor Elizabeth Harvey, to whom I am forever indebted for making this possible.

TABLE OF CONTENTS

Chapter 1. Introduction

Chapter 2. Literature and Practice Review: Typeface Legibility for Low Vision Adults

Chapter 3. Methods and Methodology: Integrating Scientific and Design Knowledge

Chapter 4. Scientific Review: The Influence of Stroke Width on Legibility

Chapter 5. Design of a Laboratory Typeface

Chapter 6. Generating Design and Interdisciplinary Knowledge

Chapter 7. Conclusions and Future Research

Appendices

LIST OF TABLES

LIST OF FIGURES

<u>Figure 26</u>: Cheltenham Old Style and Cheltenham Condensed Bold are found to be less legible (distance method) than Cheltenham Old Style Bold (Roethlein, 1912).

<u>Figure 27</u>: Memphis Bold, Memphis Medium, and Memphis Light, tested by Luckiesh and Moss (1940).

<u>Figure 28</u>: Memphis Light is found to be less legible (visibility) than Memphis Medium and Memphis Bold (Luckiesh & Moss, 1940).

<u>Figure 29</u>: Plantin and Gill Sans, Bold and Roman, tested by Shaw (1969).

<u>Figure 30</u>: Plantin Roman and Gill Sans Roman are found to be less legible (reading performance) than Plantin Bold and Gill Sans Bold, for partially sighted participants (Shaw, 1969).

<u>Figure 31</u>: Helvetica Bold and Regular, tested by Smither and Braun (1994).

<u>Figure 32</u>: Helvetica is found to be less legible (reading speed, performance errors) than Helvetica Bold for older participants (Smither & Braun, 1994).

<u>Figure 33</u>: Times Roman is found to be less legible (reading acuity, CPS) than Courier Bold for low vision participants (Mansfield et al., 1996).

<u>Figure 34</u>: Georgia Bold, Verdana Bold, Times New Roman Bold, and Arial Bold, and their Regular counterparts, tested by Sheedy et al. (2005).

<u>Figure 35</u>: Arial, Verdana, Times New Roman, and Georgia Regular are found to be less legible (visual acuity) than Arial, Verdana, Times New Roman, and Georgia Bold (Sheedy et al., 2005).

<u>Figure 36</u>: Franklin Gothic Book is found to be less legible (visual acuity) than Franklin Gothic Medium, Demi, and Heavy (Sheedy et al., 2005).

<u>Figure 37</u>: Tiresias LPfont, Arial Regular, and Freight Sans Book, tested by Beveratou (2016).

<u>Figure 38</u>: Freight Sans Book and Arial Regular are found to be less legible (reading speed) than Tiresias LPfont for partially sighted participants (Beveratou, 2016).

<u>Figure 39</u>: Maxular Rx Bold, Courier, Helvetica, and Times, tested by Xiong et al. (2018).

<u>Figure 40</u>: Helvetica and Times are found to be less legible (reading acuity, CPS) (represented by two long lines) than Maxular Rx Bold, and Helvetica is found to be less legible (CPS) (represented by one short line) than Courier, for participants with macular degeneration (Xiong et al., 2018).

<u>Figure 41</u>: Ovink Regular (2), Semi Bold (4), and Ultra Black (6), tested by Beier and Oderkerk (2019a). (Redacted)

<u>Figure 42</u>: Ovink Regular and Ovink Ultra Black are found to be less legible (letter recognition) than Ovink Semi Bold (Beier & Oderkerk, 2019a).

<u>Figure 43</u>: KBH Display Regular, KBH Text Regular, and Gill Sans Light, tested by Beier and Oderkerk (2019b).

<u>Figure 44</u>: Gill Sans Light is found to be less legible (reading acuity, CPS) than KBH Text Regular for older participants (Beier & Oderkerk, 2019b).

<u>Figure 45</u>: Visualisation of all typefaces found to have lower (grey) and higher (black) legibility included in the review.

<u>Figure 46</u>: Visualisation of all relationships between typefaces found to have lower and higher legibility included in the review.

Figure 47: The upper right quadrant contains typefaces found to have higher legibility (black ‘o’s), based on all studies included in the review.

Figure 48: The upper right quadrant contains typefaces found to have higher legibility, based on eight studies included in the review.

Figure 49: The upper right quadrant contains typefaces found to have higher legibility, based on six studies included in the review.

Figure 50: Figures 47, 48, and 49 compared.

Figure 51: Subset of scientific studies informing the laboratory typeface.

Figure 52a: Stroke width and letter width values of twenty-three sans serif text typeface families.

Figure 52b: Stroke width and letter width values of twenty-three sans serif text typeface families with axes, grid lines, and labels removed.

Figure 52c: Detail of Figure 52b.

Figure 53: Typeface width nomenclature.

Figure 54: Typeface weight nomenclature.

Figure 55: Stroke width and letter width values of the four master typefaces.

Figure 56: Stroke width and letter width values of the nine laboratory typefaces; four masters and five interpolated typefaces (see Figure 55 for parameter values of four masters).

CHAPTER 1. INTRODUCTION

1.1 INTRODUCTION

This research is motivated by the unprecedented aging of the world's populations (United Nations, 2020) and the associated increased risk of visual impairment (RNIB, 2019). The majority of people with visual impairments are not blind, having significant remaining vision known as low vision (Arditi, 2004). Difficulty with reading is a central concern for people living with low vision (Legge, 2007), suggesting an increased need for designers to produce accessible communications. Existing print guidelines regarding typographic design for low vision readers are defined by larger size recommendations, with guidance on the choice of typefaces themselves being broad, from for example the UK Association for Accessible Formats (UKAAF, 2019). Vision science research illustrates the influence of typeface characteristics on legibility (Legge, 2016), suggesting that greater specificity in typeface recommendations could increase the proportion of the population able to access printed text.

This practice-based research is based upon an initial inquiry formulated around one main question: What is the optimal typeface stroke width for low vision adults? The focus is on stroke width—also known as weight—because scientific evidence suggests that a bolder typeface can improve legibility (Arditi, 2004), however optimal values are not well understood (Bernard et al., 2013; Legge, 2016). This research is also focused on low vision legibility in the context of print, as printed (versus digital) text cannot be customised by the reader making its legibility crucial. In order to address this research question, an integration of knowledge from vision science and typographic design—the two major disciplines concerned with the topic of typeface legibility—is undertaken. I am able to perform this interdisciplinary research based on my training in both science (MSc Biology) and design (BDes Graphic Design).

The development of interdisciplinary typographic knowledge is approached through a quantitative analysis of typefaces. This approach is comprised of two main methods: the measurement of typeface proportions and information visualisation. These methods are utilised to integrate scientific knowledge generated through legibility research experiments with design knowledge residing in the form of typefaces themselves. The methods employed to determine which scientific studies to include in the analysis and which typefaces to measure are presented in the practice chapters within this thesis (Chapters 4-6). Scientific and design knowledge construction are addressed in Chapter 3 (section 3.3), with reference to design researchers including Professor Nigel Cross (2007).

The purpose of this research is to evidence and build a foundation for inclusive typographic knowledge. I define *inclusive typography* as the area of communication design focused on increasing the number of people with low vision able to access text. Inclusive typographic knowledge serves to contribute to both future legibility research and print design guidelines for low vision readers. As such, the audience for this thesis is both legibility researchers focused in this area, and communication design practitioners seeking knowledge on typographic design for low vision readers.

This PhD research began in 2007 and has therefore been undertaken across a timespan of more than a decade. The research presented within this thesis reflects the context which gave rise to the investigation, as well as the contemporary context within which it is published. The thesis demonstrates that the research question remains relevant and the need for interdisciplinary approaches to typographic research persists to this day.

1.2 KEY THEMES AND TERMINOLOGY

1.2.1 LOW VISION

Low vision compromises sight through blurring, patchiness, and loss of central or peripheral vision. In 2010, low vision was estimated to affect approximately 246 million people worldwide (Mariotti, 2012). In the United Kingdom, over two million people are estimated to be living with sight loss severe enough to have a significant impact on their daily lives (RNIB, 2018).

The main causes of sight loss in adults are age-related macular degeneration (AMD), glaucoma, cataract, and diabetic retinopathy (RNIB, 2019). Risk of sight loss increases with age, with one in nine people aged 60 years and over affected in the United Kingdom, one in five people aged 75 years and over, and one in two people aged 90 years and over (RNIB, 2018). Younger adults can also have low vision, and in this thesis *adult* is defined as aged 18 and above. The number of people in the United Kingdom with sight loss is predicted to increase dramatically over the coming decades (RNIB, 2018).

Low vision can be defined functionally as a visual impairment resulting in the inability to read the newspaper at a standard distance (40 cm) with best optical correction (i.e. prescription lenses) (Legge, 2007). Access to text is fundamental to participation in modern society, and the primary goal of vision rehabilitation is improving access to written materials (Arditi, 1996). The design community has an increasingly critical mitigating role to play in this context through the production of inclusive typographic design.

1.2.2 INCLUSIVE TYPOGRAPHY

Inclusive design is a response to the diverse demands of today’s consumers, especially those who are elderly or disabled (Clarkson et al., 2003). This increasingly pervasive approach to architectural, product and communication design, seeks to meet the needs of the largest user-group possible, whilst taking into consideration the goals of commerce (Clarkson et al., 2003). As stated in section 1.1, I define inclusive typography as the area of communication design focused on increasing the number of people with low vision able to access text.

An inability to access text cannot be solely attributed to the visual abilities of a reader. It is the congruence between visual abilities and typographic design that ultimately determines effective reading (Legge, 2007). The inclusive design community has aptly named such mismatches as “disabled by design” (Clarkson et al., 2003, p.1). Therefore, the design community has a role to play in increasing the percentage of people with low vision that can access written materials.

Inclusive typography is increasingly being adopted by practicing designers for two major reasons beyond ethical issues surrounding social equality. First, there is an economic imperative to meet the needs of the so-called ‘grey market,’ made up increasingly of the affluent and discriminating ‘baby boomer’ generation (Evamy & Roberts, 2004). Second, designers must increasingly operate within a legislative context advocating for the rights of people with disabilities. Under the Equality Act 2010, businesses or organisations in the United Kingdom are legally required to make reasonable adjustments (or changes) to avoid putting people with disabilities at a substantial disadvantage (RNIB, 2020). A key requirement of the Equality Act is the provision of accessible information, which service providers must follow (UKAAF, 2019).

1.2.3 LEGIBILITY AND LEGIBILITY RESEARCH

To date, inclusive typography guidelines are based largely on legibility research conducted within the vision science community. The term *legibility* in this discipline—and in this thesis—refers to the perceptual properties of text that influence readability (Legge, 2007). Therefore, issues of content that can render text difficult to read, do not in any way influence its legibility. Legibility of a text depends on both its local and global properties (Legge, 2007); local properties are characteristics of individual letters or groups of letters (e.g. typeface), while global properties are layout characteristics (e.g. line length). The vision science definition of legibility is distinct from that employed within the design community. The typographic design community distinguishes between *legibility* referring to the ease of recognition of letters and words, and *readability* referring to the ease and pleasantness of reading text (Felici, 2012). Note that James Felici's *The Complete Manual of Typography* (2012) is employed throughout the thesis as a standard reference for the definition of typographic terms. Within this thesis the term legibility encompasses all measures of reading performance as long as they depend on the physical properties of text (local or global). While this term may appear broad, in most circumstances legibility becomes defined more specifically according to the particular methods employed to measure it.

Legibility research within the vision sciences is conducted using what are called psychophysical methods. *Psychophysics* is the study of the relationship between physical stimuli and perceptual responses (Norton et al., 2002); in this case, the relationship between the physical properties of text (e.g. stroke width) and reading performance (Legge, 2007). The psychophysical study of reading commonly employs three legibility metrics: reading acuity, reading speed, and critical print size (CPS), which are discussed in greater detail in Chapter 2 (section 2.2.2). While legibility research is primarily conducted within the scientific community, there are increasing interdisciplinary contributions from the design research community (e.g. Bessemans, 2012; Beveratou, 2016; Dyson & Beier, 2016). Within this PhD thesis, interdisciplinary design research is included within reviews and analyses, and is referred to as scientific research if the knowledge is generated through scientific methods.

1.2.4 INCLUSIVE TYPOGRAPHY GUIDELINES

The challenge of reviewing scientific research and translating it into practical recommendations for designers is primarily undertaken by visual impairment organizations including the Royal National Institute of the Blind in the United Kingdom (RNIB, 2017), Lighthouse International in the United States (Arditi, 2018), and the Canadian National Institute for the Blind (CNIB, 2020). The UK Association for Accessible Formats, an industry association, also publishes guidelines (UKAAF, 2019) which RNIB links to from its website. Designers are also involved in this process, for example the joint publication between RNIB and the International Society of Typographic Designers (RNIB & ISTD, 2007). The print guidelines referenced above are reviewed in Chapter 2 (section 2.4) and represent a Canadian, United States, and United Kingdom perspective.

Inclusive typography guidelines do not yet rest upon a strong scientific foundation. A review of typography for readers with low vision in the *Journal of Visual Impairment and Blindness* concludes that “research has not produced consistent findings and thus that there is a need to develop standards and guidelines that are informed by evidence” (Russell-Minda et al., 2007, p.402). Criticism of guidelines is also found within empirical papers. Rubin et al. (2006, p.545) state that “the scientific basis for the guidelines is elusive at best”. Tarita-Nistor et al. (2013, p.57) remark that “no solid evidence has been provided to support these recommendations”. Hedlich et al. (2018, p.398) state that “Most recommendations addressing font styles are not evidence based.” Designers are also critical of guidelines, for example recommending that RNIB’s Clear Print guidelines “need good supporting evidence, interpreted in terms of practical document design strategies, before they become the basis for public policy” (Waller, 2011, p.11). Scientific legibility research related to low vision reading is reviewed in Chapter 2 (section 2.4). The studies included in the review are focused on a high-resolution context and utilise both print and screen-based media for legibility testing.

1.3 FOUNDATIONAL RESEARCH

This PhD builds upon a foundation of knowledge contributed by scientists and designers whose research is referenced throughout the thesis. The work of three researchers is particularly influential: Gordon E. Legge, Charles Bigelow, and Aries Arditi. The vision scientist Gordon E. Legge's research contributes crucial foundational knowledge in this subject area. Legge's book *Psychophysics of Reading in Normal and Low Vision* (2007) provides an overview of twenty seminal research papers by Legge and his colleagues on the psychophysics of reading in normal and low vision, published between 1985 to 2002. The book includes a chapter entitled "Displaying Text" which discusses the literature on the subject. More recently, Legge (2016) reviews vision science knowledge on low vision and reading, in the context of opportunities presented by digital formats. Chapter 2 (section 2.4.3.5) and Chapter 4 (section 4.3.6) examine one of the seminal research papers; *Psychophysics of Reading XV: Font Effects in Normal and Low Vision* (Mansfield et al. 1996).

Legge also collaborates with typeface designer and academic Charles Bigelow on an interdisciplinary investigation. Legge and Bigelow (2011) present evidence that the distribution of print sizes in historical and contemporary publications falls within the range of text sizes which can be read at maximum speed. Bigelow's contributions are influential for this PhD research, most notably his interdisciplinary collaborations. Chapter 2 (section 2.3.3) describes a laboratory typeface designed by Bigelow specifically to investigate the influence of serifs on legibility (Morris et al., 2002). Note that the term *laboratory typeface* is utilised within this PhD thesis, which I define as a typeface designed specifically for experimental legibility research. Chapter 4 (section 4.3.9) analyses another scientific legibility research study to which Bigelow contributes (Xiong et al. 2018).

The work of vision scientist Aries Arditi is also influential for this PhD research. Arditi is a leader in the creation of laboratory typefaces for low vision legibility research (e.g. Arditi, 2004). Chapter 2 (section 2.3.3) examines one of his research studies which employs a laboratory typeface to assess the influence of stroke width on legibility. Chapter 5 describes a laboratory typeface that I design, which builds upon the work of both Arditi and Bigelow.

1.4 PHD RESEARCH SCOPE AND TYPOGRAPHIC TERMINOLOGY

This investigation is focused on the typeface characteristic stroke width (see Figure 1), which is also known as weight. Within the design literature, *weight* is utilised to refer to “the thickness of the strokes that make up the characters of a typeface” (Felici, 2012, p.328). Within the scientific literature, this concept is referred to as *stroke width* (Legge, 2007). Within this thesis, both terms are used depending on the context. Generally, the term *stroke width* is prioritised when referring to typeface proportions and numerical values, and *weight* is used when referring to typefaces used in design practice.

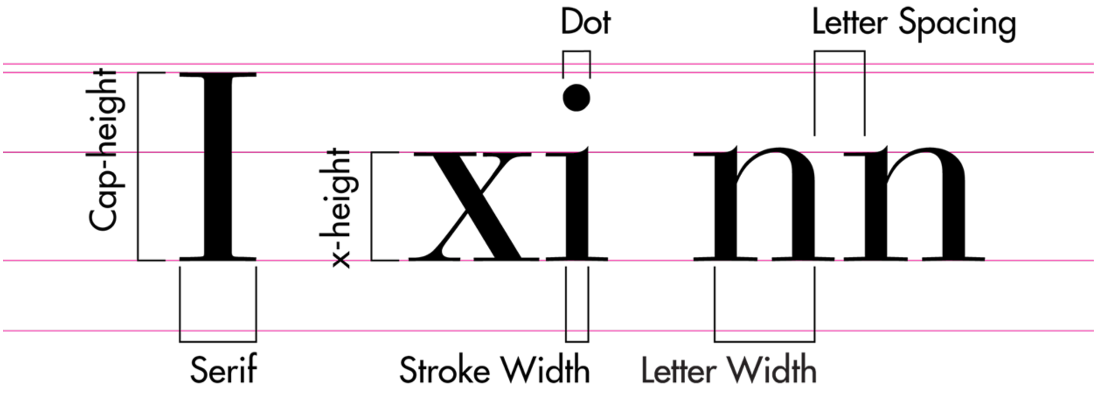


Figure 1: Typeface anatomy and letters employed for measurements, illustrated using the typeface Bodoni.

Letter width is also investigated in this thesis (see Figure 1), because this typeface characteristic varies alongside stroke width (section 2.3.1) and influences legibility (section 2.4.3.4). In the scientific literature, the term *aspect ratio* is used, referring to the width to height ratio of a character (Arditi, 1996). Within this thesis, the term *letter width* is used and refers to letter width as a proportion of letter height (i.e. synonymous with aspect ratio).

To a lesser degree, letter spacing is analysed (see Figure 1). *Letter spacing* refers to the spacing between letters as defined within a typeface (Felici, 2012). Similar to letter width, letter spacing varies alongside stroke width (section 2.3.1) and influences legibility (section 2.4.3.6). Letter spacing data is not visualised in the thesis, however it is discussed in cases when this aids in interpreting the influence of stroke width on legibility (e.g. section 4.3.6).

The investigation is primarily focused on sans serif typefaces. *Sans serif* typefaces have strokes which end in blunt terminals (Felici, 2012). *Serifs* are short lines at the ends of horizontal and vertical strokes (Cheng, 2005) (see Figure 1). Chapter 2 (section 2.4.3.2) presents evidence that sans serif typefaces are more legible for low vision adults. Chapter 4 and Chapter 6 include analyses of serif typefaces, in order to be as comprehensive as possible in the research.

The analysis focuses on text typeface families. *Text typefaces* are designed for use in long texts (Felici, 2012), with text typeface families including a range of weights and widths. A *typeface family* describes a group of typefaces that share a common root name and design characteristics (Felici, 2012), for example Helvetica Regular, Helvetica Bold, Helvetica Condensed, etc. Within this thesis, the term *text typeface* is used to mean *text typeface family*.

Lowercase versus uppercase characters are the focus of analyses, as these predominate in most English texts (Jones & Mewhort, 2004). Numerals are excluded from the investigation. *Reversed* type—white text on a black background—is also excluded from the investigation, as printed reading material is usually set black on white. Lastly, the research focuses on *Latin characters*; the characters on which Western and most Eastern European languages are based (Felici, 2012).

1.5 TYPEFACE WEIGHT: ORIGIN AND CONTEMPORARY PRACTICE

Bold typefaces have their roots in the Industrial Revolution and the birth of advertising (Haley, 2020). By the early 19th century there was a demand for display typefaces (Dodd, 2006) designed to be used at large sizes to “grab the reader’s attention” (Haley, 2020). The majority of typefounders in Britain were issuing bold display typefaces by the 1820s (Twyman, 1993). The earliest of these was known as the fat face (Twyman, 1993), the invention of which is credited to the British typefounder Robert Thorne (1754-1820) (Dodd, 2006) (Figure 2). Although the publication of *New Specimen of Printing Types, Late R. Thorne’s* dates his fat faces to 1821, it is thought that he designed the first of these in 1803 (Meggs & Purvis, 2006). Thorne took advantage of the popularity of modern typefaces (Dodd, 2006) and significantly expanded the thickness of the heavy strokes, increasing the weight and contrast (Meggs & Purvis, 2006). *Contrast* refers to the difference between the thick and thin portions of the strokes (Felici, 2012). The fat faces are described as “Bodoni or Didot designs on steroids” (Bigelow & Holmes, 2015) (see Figure 1).

Figure 2: A contemporary fat face (Solotype).

Not until the early 20th century did typefounders begin to integrate bold weights into typeface families (Bigelow & Holmes, 2015). By the late 20th century, the majority of new typeface families and revivals (e.g. Garamond) included at least two bold weights (Bigelow & Holmes, 2015). In the 21st century, new typefaces often include at least four weights (Bigelow & Holmes, 2015). For example, typeface weights for Neue Helvetica—in order from lightest to heaviest—include: Ultra Light, Thin, Light, Roman, Medium, Bold, Heavy, and Black (Linotype, 2021).

Typeface weight names are usually subjective and vary between different typefaces and languages (Bigelow, 2019). These names “give an ordinal sense of boldness” within a typeface family, with no standardisation between typeface families (Bigelow & Holmes, 2015). While this PhD research is focused on the print context, it should be noted that the World Wide Web Consortium (W3C) defines typeface weight on a numerical scale from 100 to 1000, from lightest to boldest (Bigelow, 2019). However, the scale is “imprecisely intuitive and ordinal” (Bigelow, 2019, p.166) and similar to weight names, does not describe a typeface’s numerical stroke width.

The primary use of bold typefaces in setting text is for emphasis or hierarchy (Haley, 2020). The latter refers to the creation of different levels of importance through typeface choice and text arrangement (Haley, 2020). For example, setting headings in a bold typeface is common within typographic practice (Haley, 2020). This PhD investigation demonstrates that bolder typefaces improve reading performance in the context of low vision reading. It thus encourages expanding the use of bolder typefaces within typographic practice beyond emphasis and hierarchy, in order to increase the percentage of people able to access text.

1.6 PHD RESEARCH JOURNEY

This research is conducted *through* design (Frayling, 1993), taking advantage of the unique insights gained through design practice (Godin & Zahedi, 2014). The starting points for this type of research are often issues arising from the researcher’s own practice, that can also be recognized as valid in the wider professional context (Gray & Malins, 2004). This research falls under the category of practice-based research, defined as “an original investigation undertaken in order to gain new knowledge partly by means of practice and the outcomes of that practice” (Candy, 2006, p.1). In this way, the creation of design artefacts is central to the research process, with knowledge gained through the making process, as well as being embedded within the artefacts themselves.

The impetus for this research emerged seventeen years ago. After completing my MSc in Biology in 2003, I embarked upon a new path toward becoming a graphic designer. During my first semester of design school, I woke up one morning with everything appearing darker through one of my eyes. I was diagnosed with optic neuritis—an inflammation of the optic nerve—which passed in a few weeks. The next year, I experienced central vision loss and was legally blind in one eye for weeks, during which time I was diagnosed with multiple sclerosis. Optic neuritis is a common presenting symptom of the disease, and I experienced this on a few occasions. My eyes have never been the same, and in those first years of the disease I found reading particularly difficult. I became deeply interested in communication design for people with visual impairments.

In my third year of design school, I undertook a full-year project focused on developing a (mock) visual identity for CNIB. CNIB had just published their *Clear Print Accessibility Guidelines* (CNIB, 2006), which provided my first exposure to such recommendations. I was surprised by the lack of information and specificity in the guidelines, much of which is considered simply ‘good design’. For example, the guidelines recommend “don’t crowd your text”, accompanied by a photograph of illegible text with letters touching and overlapping (CNIB, 2006, p.15). As I had a good understanding of typography, many recommendations were not useful for the development of my practice, and I remained unclear how to design for a visually impaired audience. I felt compelled to contribute to the development of inclusive typography practice.

In the final year of my undergraduate degree, I applied to the Royal College of Art (RCA) with a proposal to design a typeface for people with low vision as a research degree by project. I was inspired by the typeface Read Regular designed for people with dyslexia by Natascha Frensch (2003) at the RCA. I believed my research project was feasible based on my training in science and design, and that such a typeface would be an important contribution to the field, providing a tool for designers creating accessible communications as part of their practice. During this final year of my undergraduate degree, I began preliminary work in this area through two typeface design courses. As I familiarised myself with the scientific legibility research literature and attempted to apply it to my typeface design practice, it became clear that the scientific knowledge that would underpin a typeface for low vision did not exist.

In 2007 I began my PhD research at RCA with a new proposal; to design a laboratory typeface specifically for experimental testing. My PhD proposal directly addressed the lack of information to be found on the subject, revealed through my earlier attempts to design a typeface for low vision users. My research centres on typeface weight as this typeface characteristic had been experimentally found to influence low vision legibility (Arditi, 2004), though was not well understood. This gave rise to my main thesis research question: What is the optimal typeface stroke width for low vision adults?

My goal was to design a laboratory typeface based on both scientific and design knowledge, addressing the thesis research question experimentally. During my first year of doctoral research, I intuitively began examining typefaces like a biologist would study organisms yet doing this with the typographic knowledge of a designer. Based on my understanding of typeface anatomy, I began measuring the proportions of typefaces (section 4.2.2). I was particularly interested in measuring the typefaces that had been used as experimental test material in scientific legibility research papers, for example Franklin Gothic tested in a study by Sheedy et al. (2005). I also began measuring the proportions of text typefaces that I had become familiar with through my design training (e.g. Helvetica). I entered this data on typeface proportions into a spreadsheet and began visualising it (i.e. graphing), based on my scientific training in data analysis. Through my interdisciplinary practice (discussed in Chapter 3) and the resultant visualisations, I was finally able to make sense of the scientific legibility research literature and understand what typeface proportions were associated with improved reading performance for people with low vision. For the first time, I was also able to understand the proportions of commonly used text typefaces. This scientific and design knowledge became the basis for developing my laboratory typeface (presented in Chapter 5).

Reflecting on this research, the laboratory typeface itself was less important than what my practice had revealed. I now saw the potential to develop a foundation of interdisciplinary—science and design—typographic knowledge through information visualisation. My interdisciplinary practice had given rise to the quantitative analysis of typefaces that would become central to my PhD research. In 2009 I made a final iteration to my PhD direction and focused my research on the consolidation of scientific knowledge, generation of design knowledge, and development of interdisciplinary knowledge through information visualisation. This change in direction was also influenced by the rise of information visualisation as a creative practice occurring within graphic design (section 3.4.6.1), which has continued to the present.

Based on my scientific training, I employed information visualisation as an analytical tool (Hand, 2008) within the PhD research. For example, visualisations reveal the stroke widths of typefaces found to have higher legibility. As a design practitioner, I approached information visualisation as a communication design medium. I endeavoured to create visualisations that were not only clear in their communication, but also visually interesting and aesthetically rich. While the main audience for the visualisations is researchers and designers focused on low vision legibility, my design practice facilitated the dissemination of this research to the wider communication design community. My goal was to create visualisations that would offer a new understanding of typeface design practice, as I had experienced. My practice and information visualisation as a research method are discussed in more detail in Chapter 3 (section 3.4).

1.7 CONTRIBUTION TO KNOWLEDGE

This PhD research contributes to the area of inclusive typography through the development of interdisciplinary—scientific and design—knowledge on typeface legibility for low vision adults. Through the quantitative analysis of typefaces, scientific knowledge generated through legibility research experiments is integrated with design knowledge residing in the form of typefaces themselves. This involves the measuring of typeface proportions and the visualisation of this typeface data. The analyses focus on typeface stroke width and letter width, because these characteristics vary alongside one another.

Analysing typeface proportions may seem a natural approach to legibility research, however characterising typefaces entirely in quantitative terms is not common within either the scientific or design community. Scientific studies generally report on the relative legibility of typefaces without describing them numerically, for example reporting a low vision reading speed advantage for Courier Bold versus Times Roman (Mansfield et al., 1996). Designers similarly refer to typefaces through naming systems, for example Arial Bold 12 point, with neither weight nor point size having an accurate numerical meaning. Without numerical values associated with typefaces, questions are raised such as: how can legibility researchers develop hypotheses regarding the underlying cause for performance differences? How can legibility researchers compare results across experimental studies that test different typefaces? How can inclusive typography guidelines relate legibility research findings to commercial typefaces that may share proportions with those tested within scientific experiments?

The quantitative analysis of typefaces therefore serves to clarify and consolidate scientific research, formalise design knowledge, and facilitate the integration of knowledge across disciplines. More specifically, ten scientific studies which test the influence of stroke width on legibility are analysed. This entails measuring and analysing the stroke width and letter width of typefaces used as experimental test material. This analysis serves to elucidate the proportions of typefaces that have been found to improve reading performance based on experimental studies (i.e. scientific knowledge). This also allows for comparison across experimental studies, consolidating scientific knowledge. This consolidation of scientific knowledge addresses the thesis research question: What is the optimal typeface stroke width for low vision adults?

In order to formalise design knowledge, the stroke width and letter width of text typefaces used in design practice are measured. This is a *design phenomenology* study as defined by Nigel Cross (2007), investigating design knowledge residing in artefacts themselves (section 3.3). Typeface weight names (i.e. nomenclature) utilised in design practice are also analysed in order to assess their relationship with stroke width numerical values. This allows for an evaluation of typeface weight recommendations (e.g. “bold” for emphasis) within inclusive typography guidelines (e.g. UKAAF, 2019). Through an integration of scientific and design knowledge as proposed, this PhD research examines how the stroke width and letter width values of typefaces found to have higher legibility relate to those of text typefaces used in design practice. This addresses the research question in the context of design practice, investigating optimal typeface weights for low vision adults.

The analyses described in this section are executed through information visualisation. Information visualisation is also employed as a communication tool (Hand, 2008), facilitating the dissemination of knowledge to both legibility researchers and practicing designers. The visualisation of typeface data constitutes the central practice-based outcomes of this PhD by project. My contributions to knowledge include visualisations of:

(1) Scientific knowledge: Stroke width and letter width values of typefaces experimentally found to have higher and lower legibility;
(2) Design knowledge: Stroke width and letter width values of sans serif text typefaces used in design practice;

(3) Interdisciplinary knowledge: Relationship between stroke width and letter width values of typefaces found to have higher legibility (scientific knowledge) and typefaces utilised in design practice (design knowledge).

1.8 THESIS OVERVIEW

This thesis consists of seven chapters. Chapter 1 introduces the reader to the context—social, scientific, and design—of the investigation, the research methods, and the major outcomes and contributions to knowledge of the PhD. Chapter 2 reviews inclusive typography guidelines, the research literature on typeface legibility for low vision adults, and typefaces designed specifically for low vision reading. This chapter also presents key legibility research concepts and background, and a critical discussion of how experimental test material (i.e. the typefaces tested) impacts the application of research to design practice.

Chapter 3 describes the methodology for interdisciplinary typographic knowledge construction employed within the PhD. The chapter begins with a review of the theoretical issues regarding interdisciplinary approaches to typeface legibility. A case is made for the integration of design knowledge into scientific legibility research, and scientific knowledge into design practice. Knowledge construction within design practice is also discussed, and the investigation of 'designerly ways of knowing' through design artefacts—*design phenomenology*—is addressed (Cross, 2007). Lastly, practice-based research and the specific methods of practice employed within the PhD investigation are discussed including typeface design, typeface measurement, and information visualisation.

Chapter 4 presents a scientific review on the influence of stroke width on legibility in the context of low vision readers. Distinct from a literature review, this is a quantitative analysis of ten scientific studies. This entails the measurement and visualisation of stroke width and letter width values of typefaces experimentally found to have higher and lower legibility. Based on this consolidation of scientific knowledge, recommendations are made for inclusive typography guidelines and future legibility research.

Chapter 5 presents the design of a laboratory typeface, based on both scientific and design knowledge. The laboratory typeface is designed to test the stroke width and letter width values measured in typefaces found to have higher legibility, based on the consolidation of scientific knowledge undertaken in Chapter 4. The laboratory typeface is also based on design knowledge and reflects the stroke width and letter width values of sans serif text typefaces. The formalisation of this design knowledge is presented, specifically the measurement and visualisation of stroke width and letter width values of sans serif text typefaces used in design practice.

A rigorous approach to formalising typeface design knowledge is undertaken in Chapter 6. A points-based survey of design sources (e.g. typeface best-sellers lists) is employed to determine a group of typefaces to serve as the basis for investigation. The stroke width and letter width of twenty sans serif text typefaces are measured and visualised, and the relationship between typeface nomenclature (e.g. “bold”) and numerical values is determined. In an important culmination of the practice chapters, Chapter 6 presents the visualisation of interdisciplinary knowledge. These visualisations illustrate the relationship between the stroke width and letter width values of typefaces found to have higher legibility (scientific knowledge) and those of typefaces utilised in design practice (design knowledge).

Finally, Chapter 7 summarises the scientific, design, and interdisciplinary knowledge on typeface legibility for low vision adults contributed through the PhD investigation. The findings from each of the practice Chapters 4-6 are presented, and the contributions to knowledge are described in the context of future legibility research and the development of evidence-based inclusive typography guidelines. Chapter 7 ends by examining the limitations of this investigation and proposing future directions for legibility research.

The overall thesis structure represents my design process, and the chapters can be understood as chronological. However, because this research was conducted over a timespan of more than a decade, each chapter has been consistently updated to reflect contemporary research and practice. For example, while the scientific review (Chapter 4) was conducted before creating the laboratory typeface in 2009 (Chapter 5), I returned to the scientific review in 2020 adding four more legibility research studies.

The introductory chapter set out to provide an overview of the PhD research, the main research question, contributions to knowledge, and the context within which the research is undertaken. Next, we turn to Chapter 2 which presents a more detailed contextualisation of this research, through a literature and practice review focused on legibility and low vision. This provides the broader context through which gaps in knowledge are identified.

CHAPTER 2. LITERATURE AND PRACTICE REVIEW: TYPEFACE LEGIBILITY FOR LOW VISION ADULTS

2.1 INTRODUCTION

This chapter reviews inclusive typography guidelines, the research literature on typeface legibility and low vision, and typefaces designed for low vision reading. Through this literature and practice review, Chapter 2 demonstrates that while typeface characteristics are known to influence reading performance for low vision adults, optimal values are not well understood. My research responds to this gap in knowledge through the research question: What is the optimal typeface stroke width for low vision adults?

This chapter begins with an introduction to the historical and contemporary context of legibility research, addressing underlying factors which have contributed to a lack of knowledge in this subject area. Commonly employed legibility metrics (e.g. reading speed) and the three major categories of reading deficits experienced by low vision adults are presented. This is followed by an examination of legibility research methodology, with a focus on experimental test material (i.e. the typefaces tested) and how this can limit the application of scientific research to design practice. This background information contextualises the literature review that follows.

The literature review addresses the influence of character size on low vision reading and the influence of typeface characteristics. Stroke width is addressed, as this is the focus of the PhD research. Because typeface stroke width varies alongside letter width and letter spacing, the review also addresses these characteristics. The relative legibility of serif and sans serif typefaces is also considered, as the PhD research focuses on sans serif typefaces. Monospaced versus proportionally spaced typefaces are also discussed, as evidence suggests that monospaced typefaces improve reading performance for people with low vision. The practice review follows and presents five typefaces designed specifically for low vision reading and associated legibility testing.

The inclusive typography guidelines reviewed apply to a print context, while the experimental studies presented apply to a high-resolution context and utilise both print and screen-based media for legibility testing. While the majority of legibility research is conducted within the scientific community, design research is also included in the literature review (Bessemans, 2012; Bessemans, 2016a; Beveratou, 2016; Beier & Oderkerk, 2019a). Ten of the studies introduced in this chapter are analysed in detail within a scientific review in Chapter 4.

2.2 KEY CONCEPTS AND BACKGROUND ON LEGIBILITY RESEARCH

2.2.1 LEGIBILITY RESEARCH HISTORY AND CONTEMPORARY CONTEXT

Historically, typeface legibility has not been a priority research interest within the vision sciences. Real-world tasks such as reading have often been ignored or studied within a context of higher levels of cognitive processing (e.g. learning to read) (Legge, 2007). While there was a temporary flurry of legibility research in the 1960s and 70s (Lund, 1999), the work typically focused on people with normal sight versus low vision (Russell-Minda et al., 2007). Since Miles Tinker's influential book *Legibility of Print* (Tinker, 1963), relatively few studies on the influence of typeface on reading have been published (Legge, 2007). A contemporary review of reading research similarly assesses that the main area of focus has moved away from typography to a focus on cognitive mental processes (Beier, 2016). While design researchers are increasingly undertaking interdisciplinary legibility research (section 3.2.1), there are still relatively few researchers who have designers as their main audience and this area is "yet to make the big break within design research" (Beier, 2016, p.65).

Of the vision scientists employing psychophysical methods to study low vision reading, the work of Gordon E. Legge is notable. Legge states in his book *Psychophysics of Reading in Normal and Low Vision* (2007) that much of the research in this area is focused on characterising reading deficits, with applications to displaying text being secondary. Furthermore, research focused on typeface legibility often tests normally sighted people near the acuity limit (acuity discussed in section 2.2.2), versus low vision research participants. Lastly, due to challenges surrounding the controlled investigation of typefaces—which differ across characteristics simultaneously—there can be limitations in the application of scientific research to design practice. This issue is addressed in section 2.3.

2.2.2 LEGIBILITY RESEARCH METRICS

As stated previously (section 1.2.3), the psychophysical study of reading commonly employs three legibility metrics: reading acuity, reading speed, and critical print size. Reading acuity is the minimum size, or equivalently the greatest distance, at which text can be read (Lovie-Kitchen & Whittaker, 1998). Reading acuity is measured using words or sentences, while letter acuity—a related metric—is measured using short strings of unrelated test letters (Legge, 2007). Reading acuity is a more functionally relevant measure of vision than letter acuity, because it “includes the effects of the cognitive and visual factors that are involved in a normal reading task” (Legge, 2007, p.171). This includes the effects of context “in which the reader may be able to determine letter and word identities based on the context of the other words that have already been read in the sentence” and the effects of ‘crowding’ of nearby letters and words “which may make the reading task more difficult” (Legge, 2007, p.171) (crowding discussed in section 2.4.3.6). “Reading acuity is highly correlated with letter acuity.” (Legge, 2007, p.50).

One criticism of acuity-based metrics is the assumption that better acuity will also result in increased legibility at larger sizes (Legge, 2007). Evidence supporting this critique is provided by Mansfield et al. (1996) who investigate the legibility of two typefaces; Courier Bold and Times Roman. For normally sighted subjects, Courier Bold is found to be more legible based on reading acuity, however maximum reading speeds are faster with Times Roman.

Arditi argues that reading acuity is an appropriate legibility metric within a low vision context (Arditi, 1996), as people with low vision often read at their acuity limit (Arditi, 1996; Legge, 2007). One further criticism of acuity-based metrics is that they cannot be used to assess the performance of typefaces at different sizes (Legge, 2007). Lastly, acuity-based metrics are not well suited to investigating the global properties of text (Legge, 2007).

Reading speed (words per minute) is another common legibility metric. Test material utilised in such experiments ranges from single sentences to multiple paragraphs. In contrast to reading acuity, reading speed tests are well suited to investigating the influence of text size on reading, as well as the global properties of text (Legge, 2007). Researchers who prefer this method (e.g. Tinker, 1963; Legge, 2007) do so because it has more ecological validity (i.e. more closely resembles natural reading outside of the laboratory) and is functionally significant for people with low vision who commonly read at a slow pace. A reading speed test may involve recording the number of words or letters read within a specified timeframe (e.g. Beveratou, 2016) or recording the time it takes for a participant to complete a reading task (e.g. Smither & Braun, 1994). Maximum reading speed (MRS) is a specific metric, representing “the best reading performance that can be attained when print size is not a limiting factor” (Legge, 2007, p.171).

Critical Print Size (CPS) is another important metric. CPS is a measure of the smallest type size that supports the maximum reading speed. Experiments find that reading speed increases sharply with character size up to the CPS at which point the relationship plateaus, and ultimately at very large letter sizes reading speed slowly declines (Legge, 2007).

Rapid serial visual presentation (RSVP) is another method used to measure reading speed. RSVP presents words rapidly at a fixed location on a video display, removing the need for eye-movements in reading. RSVP allows for significantly faster reading speeds up to three to four times faster than normal (Rubin et al., 1992).

Assessing legibility based on reading performance has limitations. This includes individual differences in reading ability due to nonvisual factors (e.g. education), performance differences due to contextual factors, and the challenge of sourcing appropriate text for legibility testing (Legge (2007). However, methods have been developed which have proven to be sensitive to visual factors. Notably, the MNREAD Acuity Chart is a research instrument that is used to assess how a person’s reading performance is affected by print size (Mansfield et al., 1993). It is a continuous text reading acuity chart which presents a series of sentences at progressively smaller sizes. The sentences contain the same number of characters and have the same spatial layout, therefore differences in reading performance can primarily be attributed to differences in print size. Reading speed data collected using the MNREAD Acuity Chart reveal three functional measures of reading performance: reading acuity, MRS, and CPS. Versions of the MNREAD Acuity Chart have been used to compare reading performance with different typefaces (Mansfield et al., 1996; Xiong et al., 2018). The literature review in section 2.4 presents legibility research studies which employ the methods introduced in this section.

2.2.3 LEGIBILITY AND LOW VISION

While legibility depends on the physical properties of text, it is ultimately determined by a person’s visual processing capabilities (Legge, 2007). Low vision reading difficulties are commonly classified under three major categories: low acuity, reduced contrast sensitivity, and visual field loss (Legge, 2007). People with low vision typically have a reduced acuity reserve, meaning that the critical text size at which reading begins to deteriorate occurs at much larger sizes than for normally sighted people (Yager et al., 1998). As stated in section 2.2.2, it is common for low vision reading to occur at the acuity limit (Arditi, 1996). Magnifying text is the most common and effective method for increasing people’s ability to identify visual patterns, and hence improving low vision reading (Arditi, 1996). A reduced contrast reserve is also common for people with low vision, meaning similarly that the critical contrast at which reading performance declines occurs at much higher contrasts than normal (Legge, 2007). Contrast can be understood through the example of text at maximum contrast which achieves the largest possible luminance difference between black type and a bright background (Legge, 2007).

Visual field loss is also common for people with low vision, particularly central-field loss caused by AMD, which is the primary cause of visual impairment in developed countries (Legge, 2007). The functional impact on reading is a reduction in the size of the *visual span*, the number of letters that can be reliably recognised without eye movement (Legge, 2007). Normal reading involves a series of fixations on a line of text, separated by saccadic eye movements. “The size of the visual span limits the number of letters that can be recognized in each fixation.” (Legge, 2007, p.68). Reading speed is strongly correlated with the size of the visual span (Legge, 2007). Good reading performance is ultimately achieved through congruence between the visual processing abilities of the reader and the physical properties of text.

2.3 LEGIBILTY RESEARCH METHODOLOGY

2.3.1 INTRODUCTION

This section focuses on an important methodological issue which contributes to a lack of knowledge on typeface legibility. Specifically, the majority of legibility research studies test unmodified commercial typefaces, which differ across characteristics simultaneously. While such studies may be able to conclude that one typeface is more legible than another, the underlying reason for the performance difference is not clear. This has led to laboratory typefaces created specifically for the controlled investigation of typeface characteristics on legibility. However, laboratory typefaces that are unconventional in their construction—as compared to commercial typefaces—limit the application of research to design practice.

The relationship between stroke width, letter width, and letter spacing is particularly relevant to this PhD research. Within a typeface family, a bold typeface generally has an increased letter width and decreased letter spacing, compared to the regular weight (Figure 3). Stroke width, letter width, and letter spacing all influence legibility, demonstrated in the literature review (section 2.4). This highlights the challenges surrounding the controlled investigation of typefaces.

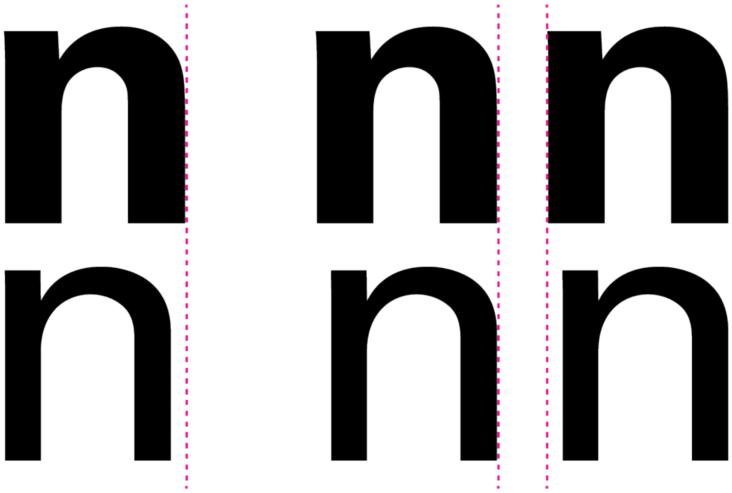


Figure 3: Lowercase ‘n’ of Univers Bold (top) and Regular (bottom). Bolder typefaces generally have increased letter width and decreased letter spacing.

2.3.2 COMMERCIAL TYPEFACES AS TEST MATERIAL

In the pre-digital era, manipulating typeface parameters to study their influence on legibility was difficult and expensive. This contributed to the common practice of testing unmodified commercial typefaces that differ simultaneously across parameters, ultimately resulting in a paucity of controlled investigation. As long ago as 1948 Curt Berger wrote: “the considerable amount of experimental work done with printing type deals almost exclusively with type already developed and not with an analysis of those factors which are the elements of all type”, specifically referencing stroke width and letter width (Berger, 1948, p.517). The lack of controlled investigation within legibility research is criticised within the design (von Ompteda, 2009a; Beier & Dyson, 2014; Bessemans, 2016a) and scientific (Arditi & Cho, 2005) communities.

The study by Sheedy et al. (2005) offers an example of the use of commercial typefaces as experimental test material to study stroke width. Sheedy et al. (2005) test four weights of the typeface Franklin Gothic (Book, Medium, Demi, and Heavy) with normally sighted participants (Figure 4). The test material is presented as uppercase letters, lowercase letters, and lowercase words. Measuring legibility as visual acuity, the experiment finds the lightest weight typeface (Franklin Gothic Book) to be less legible than the three heavier weight typefaces (Franklin Gothic Medium, Demi, and Heavy). Because these typefaces differ in more than one characteristic, it is not clear what to attribute the performance difference to. The authors acknowledge this stating: “Bold letters have wider stroke widths, but the entire character is also wider; increased legibility could be attributable to one or both factors.” (Sheedy et al., 2005, p.803). This study and the differences across Franklin Gothic typeface weights are analysed in Chapter 4’s scientific review (section 4.3.7).

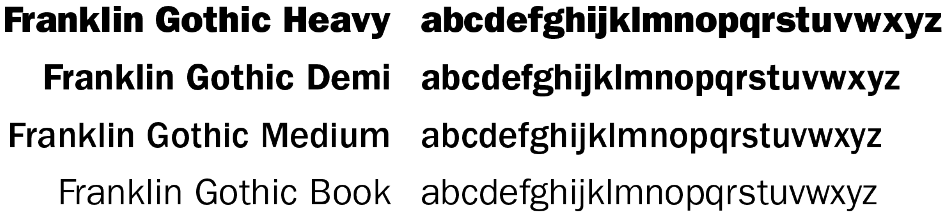


Figure 4: Franklin Gothic Heavy, Demi, Medium, and Book (top to bottom), tested by Sheedy et al. (2005).

2.3.3 LABORATORY TYPEFACES AS TEST MATERIAL

Scientists have long been interested in the controlled investigation of typefaces; the systematic modification of parameters in order to investigate their influence on legibility while holding others constant. The work of Curt Berger offers an early example of this type of work. For one of his experiments, Berger (1948) created ten figures of different widths and equal heights (five 0's and five 5's). The experiment found a general increase in legibility (measured as acuity) with an increase in numeral width (Berger, 1948). Due to technical limitations and high costs involved in the process, experiments like this were small in scale (Arditi, 1996).

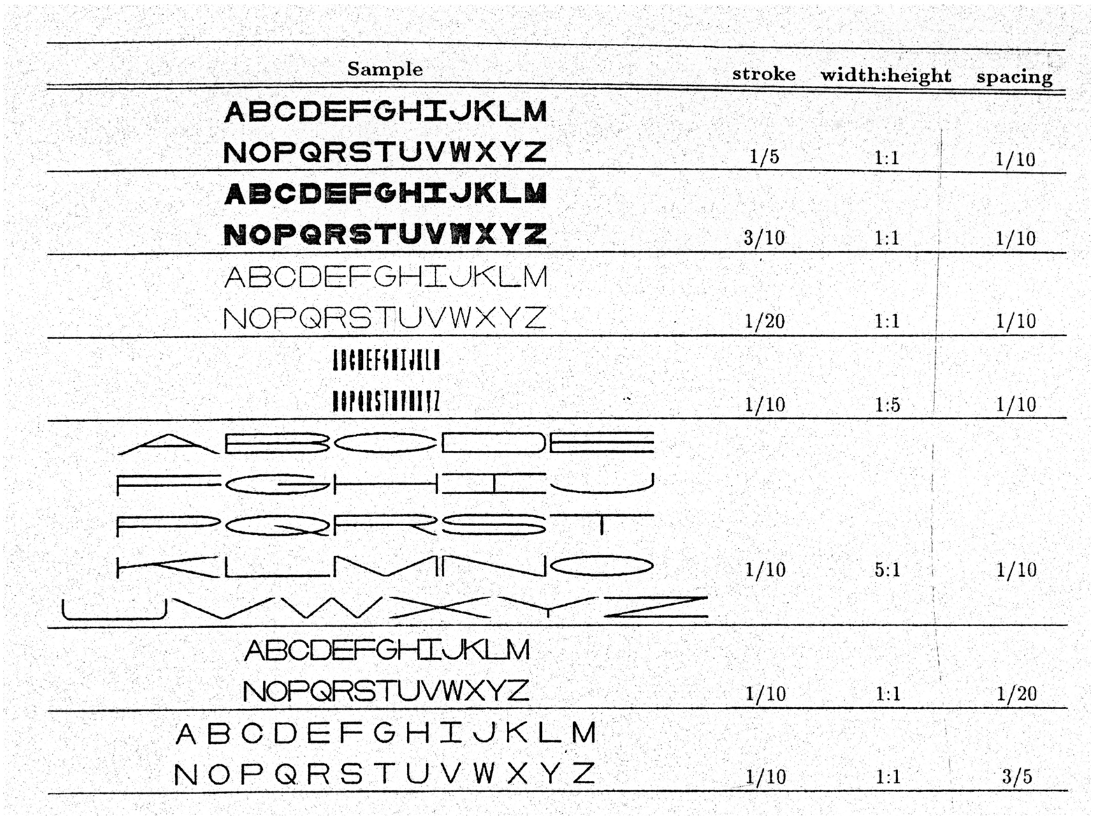
Only since the advent of digital typography has it become truly feasible to modify typeface parameters for experimental purposes. Relatively recently, legibility research has employed so-called parametric fonts as test material. Originally developed by the computer scientist Donald Knuth, the characteristics of parametric fonts can be mathematically altered in order to produce an infinite variety of fonts (Knuth, 1986). Adapted to a legibility context, these laboratory typefaces can be adjusted systematically to allow for controlled investigation into the influence of specific parameters on reading performance. The research community has commented on the value of such a methodology for legibility research (Russell-Minda et al., 2007), yet experimental literature in this area is still relatively rare.

The work of vision scientist Aries Arditi (and colleagues) is an exception, having made the most notable contribution to this method (Arditi et al. 1995a; Arditi et al. 1995b; Arditi, 2004; Arditi & Cho, 2005). Arditi has investigated the influence of typeface parameters on legibility, through systematic and carefully controlled study made possible through custom laboratory typefaces. While Arditi has read a notable amount on the subject of typeface design, as well as involving a typeface designer in some of his research projects, he has not asked a professional typeface designer to develop his test material for him, and took this task on himself (von Ompteda, 2009a). Compared to typefaces designed for continuous reading, Arditi's laboratory typefaces are unconventional in their construction (von Ompteda, 2009a). When experimental test material differs from the conventions of typeface design, there can be limitations in applying the research to design practice (von Ompteda, 2009a). The importance of incorporating design knowledge into legibility research test material is often addressed by interdisciplinary design researchers (e.g. Beier & Dyson, 2014; Bessemans, 2016a).

Arditi et al. (1995a) offers an example of a laboratory typeface designed for the controlled investigation of stroke width and letter spacing with normally sighted participants (see Figure 5). This is the only published example of a laboratory typeface created to test the isolated influence of stroke width on legibility. Arditi's typeface is constructed using the METAFONT computer language and is based on the Sloan optotypes (ten uppercase characters used for clinical acuity testing in the United States). Arditi et al. (1995a) present five-letter text strings and measure legibility as letter acuity. Five different stroke widths and three letter spacings are tested, while holding letter width constant. The five stroke widths tested are: 30%, 20%, 10%, 5%, and 2.5% (percentage of cap height). The three letter spacings tested are: 40%, 10%, and 2.5% (percentage of cap height). *Cap height* refers to the height of the capital letters of a typeface (Felici, 2012) (Figure 1). See Figure 5 for a subset of typefaces tested by Arditi et al. (1995a).

Arditi et al. (1995a) find that for the widest spaced typefaces, the thinnest and thickest stroke width variants are least legible, and the intermediate weights are most legible. Specifically, legibility decreases with a stroke width measuring 30% of cap height. Note that only the results for the widest spaced letterforms (40% cap height) are presented here, as the legibility of stroke width can be considered with the least influence of crowding, an issue discussed in section 2.4.3.6.

While Arditi's laboratory typeface is capable of controlled investigation, the high level of control also results in some of the test material deviating from the conventions of typeface design. This is particularly evident in the heaviest weight typeface (30% cap height), where the white internal spaces of many letters (e.g. W) are ‘filling in’, compromising legibility (von Ompteda, 2009a) (Figure 5). In conventional typeface design practice, increases in stroke width are accompanied by increases in letter width, in order to maintain the letter counter. *Counters* are the open, negative space inside certain characters (Felici, 2012). The letterform distortions within Arditi’s laboratory typeface limit the application of this research to design practice.



Sample	stroke	width:height	spacing
ABCDEFGHIJKLM NOPQRSTUVWXYZ	1/5	1:1	1/10
ABCDEFGHIJKLM NOPQRSTUVWXYZ	3/10	1:1	1/10
ABCDEFGHIJKLM NOPQRSTUVWXYZ	1/20	1:1	1/10
ABCDEFGHIJKLM NOPQRSTUVWXYZ	1/10	1:5	1/10
ABCDE FGHIJ PQRST KLMNO UVWXYZ	1/10	5:1	1/10
ABCDEFGHIJKLM NOPQRSTUVWXYZ	1/10	1:1	1/20
ABCDEFGHIJKLM NOPQRSTUVWXYZ	1/10	1:1	3/5

Figure 5: Laboratory typefaces designed to investigate the influence of stroke width, letter width, and letter spacing on legibility (Arditi, 1996). The top three typefaces’ stroke widths are 20%, 30%, 5% (top to bottom) with letter spacing at 10%. Used with permission of Elsevier Science & Technology Journals, from *Remediation and Management of Low Vision*, Typography, Print Legibility and Low Vision, Arditi, A., pp.237-248, Copyright (1996); permission conveyed through Copyright Clearance Center, Inc.

Experimental control does not have to come at the cost of application to design practice. The laboratory typeface designed by Charles Bigelow for the controlled study of serifs on reading performance offers a good example of this (Morris et al., 2002) (Figure 6). As serif and sans serif typefaces differ across parameters beyond the serifs themselves, Bigelow designed an intermediate style of typeface differing only in the absence or presence of serifs. Testing normally sighted participants using RSVP, Morris et al. (2002) find that the sans serif typeface is read approximately 20% faster at a small size.

Lucida Sans RSVP abcdefghijklmnopqrstuvwxyz
Lucida RSVP abcdefghijklmnopqrstuvwxyz

Figure 6: Lucida Sans RSVP and Lucida RSVP designed by Charles Bigelow to investigate the influence of serifs on legibility (Morris et al., 2002). *Society for Information Display International Symposium Digest of Technical Papers*, 33, Morris, R., Aquilante, K., Yager, D. and Bigelow, C., Serifs Slow RSVP Reading at Very Small Sizes, but Don't Matter at Larger Sizes, pp.1-4, 2002. Courtesy of Charles Bigelow © 2002.

The typographic sophistication employed in this study results in control over typeface parameters, while remaining applicable to design practice. Morris et al. (2002) illustrate the potential of interdisciplinary approaches and the importance of embedding design knowledge within scientific legibility research (addressed in section 3.2.1). Issues surrounding the controlled investigation of typefaces are addressed throughout the following literature review.

2.4 LITERATURE REVIEW: TYPEFACE LEGIBILITY AND LOW VISION

2.4.1 INTRODUCTION

The following sections review inclusive typography guidelines and scientific legibility research with a focus on low vision reading. This explores the advice given to designers regarding communications for people with low vision, and how this relates to scientific knowledge on typeface legibility. Five print design guidelines introduced in Chapter 1 (section 1.2.4) are considered, alongside eighteen empirical studies. The influence of both typeface size and typeface characteristics on legibility are addressed. This PhD research contributes to the development of evidence-based inclusive print guidelines, therefore reviewing both guidelines and scientific research is necessary to establish a gap in knowledge.

My review does not include the signage literature. This is in keeping with other reviews on low vision reading (Russell-Minda et al., 2007; Legge, 2016) and empirical studies testing the influence of boldness on legibility (Bernard et al. 2013; Beier & Oderkerk, 2019a). While similar principles apply to typeface legibility in print and signage—for example the avoidance of crowding (Beier, 2016)—differences exist between these reading contexts including reversed type, reading substrates (e.g. reflective materials), and ambient conditions (e.g. night-time viewing). As such, optimal stroke width considerations can differ between printed reading material and signage. For example, most highway signs present white letters on a coloured background (reversed type) (Legge, 2007). Their legibility can be compromised by an effect know as irradiation (or halation), in which the white strokes of letterforms appear to bleed into the surrounding background (Legge, 2007). Highly reflective signs appear to increase this effect (Legge, 2007). A typeface designed for highway signs called Clearview reduces irradiation through a decreased stroke width, and is found to be more legible than the standard U.S. highway typeface in night-time conditions (Garvey et al., 1998). The optimal stroke width for Clearview is therefore determined by different considerations than typefaces for print.

2.4.2 THE INFLUENCE OF TYPEFACE SIZE ON LEGIBILITY

Character size is a critical typographic variable for low vision reading (Legge, 2007) and inclusive typography guidelines are defined primarily by size recommendations (RNIB & ISTD, 2007). Most books and magazines are set between 8 to 10 point (RNIB & ISTD, 2007) (Figure 7). Print size in everyday text is too small, resulting in reading challenges for almost everyone with low vision (Legge, 2016). UKAAF (2019) recommends a 12 point minimum and ideally 14 point for Clear Print documents which are designed for general use, in order to reach wider audiences. According to Hugh Huddy (previously employed at RNIB), Clear Print addresses “the proportionally larger group who have sight loss but who aren’t registered as partially sighted or blind”, who “wouldn’t say they need 16 point but might say they struggle to read anything smaller than 12 point” (Waller, 2011, p.19). In the UK, these guidelines have been widely adopted in the public sector (Waller, 2011). UKAAF (2019) recommends a 16 point minimum and ideally 18 point for the more specialised Large Print format (Figure 7). Large Print is an alternative format, for example a bank providing Large Print statements to visually impaired customers (e.g. HSBC UK, 2022). UKAAF (2019) give precise x-height recommendations of 2mm for 12 point, 2.3mm for 14 point, 2.8mm for 16 point, and 2.9mm for 18 point. A typeface’s *x-height* is a measure of the height of the lowercase characters (Felici, 2012), and is based on the height of the lowercase ‘x’. An empirical investigation concludes that setting text at 16 instead of 10 point would increase the proportion of the population able to read fluently (>85 words per minute) from 88.0% to 94.4% (Rubin et al., 2006).

Helvetica 8 point
Helvetica 10 point
Helvetica 14 point
Helvetica 16 point
Helvetica 18 point

Figure 7: Helvetica set at 8, 10, 14, 16, and 18 point.

There are constraints on large character size, which are economic, technology-based, as well as functional. A critical issue for print designers is that increased character size translates into larger, heavier, and more costly documents (RNIB & ISTD, 2007). While digital formats bring unprecedented opportunities to customise text for low vision reading, the small size of screens on mobile devices also pose limitations on character size (Legge, 2016). Lastly, there is a functional constraint on character size in the context of low vision reading. Vision rehabilitation works on the principle that text should be magnified as little as possible to reach a person’s CPS. This is because there is often a trade-off between text magnification and the number of letters that fit into the field of view. As introduced in section 2.2.3, when fewer letters can be recognised without eye movements, reading performance is impacted (Legge, 2007).

2.4.3 THE INFLUENCE OF TYPEFACE CHARACTERISTICS ON LEGIBILITY

2.4.3.1 INTRODUCTION

While specific point sizes are recommended for Clear and Large Print formats, advice on typeface choice is generally broad. Vision science research illustrates the influence of typeface characteristics on low vision legibility (Legge, 2016), suggesting that greater specificity in typeface recommendations could increase the percentage of the population able to access text. For people with low vision who often read at their acuity limit, a highly legible typeface can make the difference between being able to access text or not (Arditi, 2004).

Traditionally, research into the relative legibility of commonly used typefaces found only negligible differences (e.g. Tinker, 1963), however the majority of this work was done under optimal conditions (Morris et al., 2002). Contemporary vision science research finds significant differences between the legibility of typefaces for both normally sighted people and those with low vision, reading at threshold type sizes (e.g. Xiong, et al., 2018). A study by Arditi (2004) using prototype software (Font Tailor) offers low vision participants an opportunity to adjust typeface parameters in order to meet their visual needs (Figure 8). The study finds legibility enhancements through adjustments of x-height, serif size, stroke width, letter width, and letter spacing. Measuring reading acuity, the total legibility gain from all adjustments averages 75%.

The next sections present inclusive typography guidelines regarding typeface choice alongside legibility research findings. The review addresses typeface characteristics including serifs, stroke width, letter width, and letter spacing (Figure 1). As part of the review, monospaced versus proportionally spaced typefaces are introduced and discussed.

Lorem
ipsum dolor
sit amet,

Figure 8: The average typeface created in Font Tailor (Arditi, 2004). This typeface represents “the average (arithmetic mean) of all parameters adjusted” by low vision participants using prototype software (Arditi, 2004, p.478). Reprinted from *Ergonomics*, 47(5), Arditi, A., Adjustable Typography: An Approach to Enhancing Low Vision Text Accessibility, pp.469-482, Copyright (2004), with permission from Taylor & Francis.

2.4.3.2 SERIF AND SANS SERIF TYPEFACES

Inclusive typography guidelines recommend serif and sans serif text typefaces, which ultimately encompasses a wide range of typeface characteristics. The joint publication between RNIB and ISTD (2007) *Inclusive Design: Clear and Large Print Best Practice Guide for Designers* advises that “Most typefaces can be used for Clear and Large Print, providing they are used at a reasonable size and weight.” (RNIB & ISTD, 2007, p.39). Lighthouse International’s *Making Text Legible: Designing for People with Partial Sight* also suggests “standard” serif or sans serif typefaces “with familiar, easily recognizable characters” (Arditi, 2018). CNIB’s *Clear Print Accessibility Guidelines* (2020, p.12) recommend “standard” typefaces, offering Arial and Verdana as “excellent choices”. RNIB’s *Top Tips for Creating Accessible Print Documents* are more specific and recommend a “plain sans serif” typeface, offering the examples of Arial and Helvetica (RNIB 2017, p.1). The UK Association for Accessible Formats’ (UKAAF) *Creating Clear Print and Large Print Documents* advises a “Legible, sans serif, typeface such as Arial” (UKAAF, 2019, p.3). The guidelines also refer to a lack of definitive evidence regarding the relative legibility of serif and sans serif typefaces, and list “examples of legible typefaces” including Arial, Verdana, Trebuchet, and Times New Roman (UKAAF, 2019, p.10).

Evidence suggests that readers with visual impairments prefer (Campbell et al., 2005; Russell-Minda et al., 2007; Hedlich et al., 2018) and may read better (Shaw, 1969; Morris et al., 2002; Arditi & Cho, 2005) with sans serif typefaces. Shaw (1969) investigates the legibility of a sans serif (Gill Sans) and serif (Plantin) typeface with partially sighted participants (Figure 9). The study finds that the sans serif typeface is slightly more legible. Specifically, improved reading performance with Gill Sans is measured in participants with macular degeneration. Shaw (1969) notes that the effect of typeface (i.e. serif versus sans serif) is of minor importance compared to the effect of size and weight. Shaw’s (1969) investigation of typeface weight is included in Chapter 4’s scientific review (section 4.3.4).

Gill Sans abcdefghijklmnopqrstuvwxyz
Plantin abcdefghijklmnopqrstuvwxyz

Figure 9: Gill Sans and Plantin, tested by Shaw (1969). See section 4.3.4 for Shaw’s (1969) weight investigation and all typefaces tested.

Laboratory typefaces are also utilised to examine the influence of serifs on legibility. As introduced in section 2.3.3, the study by Morris et al. (2002) tests a design-informed laboratory typeface with normally sighted participants (Figure 6). Morris et al. (2002) find that the sans serif typeface is read approximately 20% faster at a small size (~4 point), with this advantage disappearing at a larger size (~16 point). The authors conclude that serifs interfere with reading at very small sizes.

Vision scientists Arditi and Cho (2005) use a laboratory typeface to study the influence of serifs on legibility (Figure 10). This study finds a slight legibility benefit for serif typefaces when the text is small or distant, however the authors attribute this to the concomitant increase in spacing (letter spacing is discussed in section 2.4.3.6). Further, the predicted increase in legibility due to increased spacing is less than that observed, therefore the authors conclude that at very small letter sizes close to the acuity limit, serifs may actually interfere slightly with legibility.

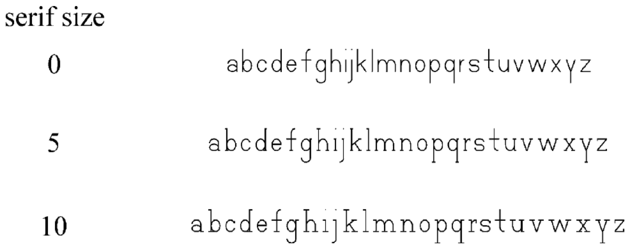
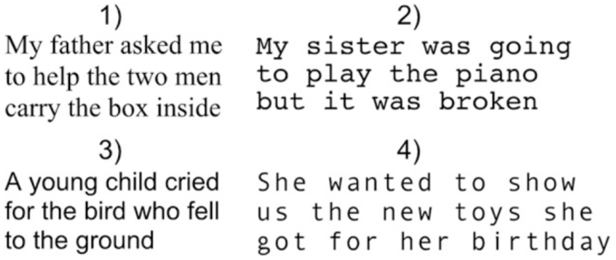


Figure 10: Laboratory typeface designed to test the influence of serifs on legibility (Arditi & Cho, 2005). Reprinted from *Vision Research*, 45(23), Arditi, A. and Cho, J., Serifs and Font Legibility, pp.2926-2933, Copyright (2005), with permission from Elsevier.

Tarita-Nistor et al. (2013) test four typefaces (Times New Roman, Courier, Arial, and a modified version of Andale Mono) with subjects with AMD (Figure 11). The study finds that near the acuity limit, performance is significantly worse with Arial than Times New Roman, Courier, and Andale Mono. The study also finds that near the acuity limit, performance is significantly better with Courier than Times New Roman, Arial, and Andale

Mono. Legge (2016) comments on the surprising result regarding Arial, noting that “The prevailing opinion has been that sans serif fonts are slightly more legible than serif fonts for low vision.” (Legge, 2016, p.8). As this test material does not control parameters, it is not clear whether the performance differences are due to serifs themselves. For example, based on my calculations undertaken in Chapter 4, Times New Roman has notably larger letter spacing than Arial (section 4.3.7). Based on the overall evidence regarding the legibility of serif and sans serif typefaces for low vision adults, this PhD investigation emphasises the analysis of sans serif typefaces.



1)
My father asked me
to help the two men
carry the box inside

2)
My sister was going
to play the piano
but it was broken

3)
A young child cried
for the bird who fell
to the ground

4)
She wanted to show
us the new toys she
got for her birthday

Figure 11: Times New Roman (1), Courier (2), Arial (3), and a modified version of Andale Mono (4), tested by Tarita-Nistor et al. (2013). This article was published in *Canadian Journal of Ophthalmology*, 48(1), Tarita-Nistor, L., Lam, D., Brent, M., Steinbach, M. and Gonzalez, E., Courier: A Better Font for Reading With Age-Related Macular Degeneration, pp.56-62, Copyright Elsevier (2013).

2.4.3.3 STROKE WIDTH

This PhD investigation is focused on stroke width; a fundamental characteristic relevant to both serif and sans serif typefaces. Inclusive typography guidelines recommend a range of weights, from regular to bold, ultimately recommending typefaces with a wide variety of characteristics. RNIB and ISTD’s guidelines (2007, p.33) discourage the use of light weight typefaces and recommend “a heavier weight of type; for example, roman, semibold, and bold”. The guidelines suggest to “Consider using a roman weight of a font, with the heavier weight used for headings and titles.” (RNIB & ISTD, 2007, p.33). They advise that “using bold text throughout the document might make it difficult and monotonous to read” (RNIB & ISTD, 2007, p.33). The guidelines also discourage the use of typefaces with extreme contrasts of weight and give Bodoni as an example. CNIB similarly discourages light weights and recommends typefaces with “medium heaviness”, and the use of “bold or heavy” typefaces for emphasis (CNIB, 2020, p.14). RNIB (2017) and UKAAF (2019) also recommend the use of bold typefaces for emphasis. RNIB (2017, p.1) recommends to “Use bold text sparingly for emphasis.” UKAAF (2019, p.9) state that “In general, bold text is preferred as a method of emphasizing text.” The guidelines acknowledge that a user may require “ordinary body text to be bold” and in this case, an alternative method must be used for emphasis (UKAAF, 2019, p.9). UKAAF’s guidelines address the topic of light weights stating that “Some typefaces have inaccessible forms, for example Arial comes in a light version which is not accessible.” (UKAAF, 2019, p.10). Lighthouse International’s recommendations (Arditi, 2018) do not address the subject of weight.

Optimum stroke width values are not well understood, reflected in the range of weights recommended within inclusive typography guidelines, and the advice to set continuous text in regular weights (RNIB & ISTD, 2007) and use bold for emphasis (RNIB, 2017), which is essentially standard typographic practice (see section 1.5). Bernard et al. (2013) corroborate this, stating that little is known regarding the influence of stroke width on reading. Legge (2016) further states that the influence of stroke width on low vision reading remains to be determined empirically.

Stroke width being a typeface characteristic worthy of further investigation, lies at the foundation of this PhD research. This informed stance is supported by other researchers. Legge (2016, p.8) states that “It is widely held that bold print is desirable for low vision.” Bernard et al. (2013, p.33) further state that “People with central vision loss often prefer boldface print over normal print for reading.”

Experimentally, increased stroke width is shown to improve legibility for people with low vision (Shaw, 1969; Arditi, 2004), people with normal vision at threshold sizes (Roethlein, 1912; Arditi et al., 1995a; Sheedy et al., 2005) and small visual angle (Beier & Oderkerk, 2019a), as well as older readers (Smither & Braun, 1994). Yet stroke thickness comes at the cost of decreased counters. Stroke widths of intermediate values are understood to improve legibility, while stroke widths that are too small or large result in legibility declines (Bernard, et al., 2013; Beier & Oderkerk, 2019a). While increases in stroke width are generally accompanied by increases in letter width to maintain counters (section 2.3.3), bolder typefaces still have smaller counters than lighter weights within the same typeface family (see Figures 3 and 4). Chapter 4 consolidates scientific knowledge on the influence of stroke width on legibility through an analysis of ten studies which test commercial typefaces. Five of these studies are referenced in this section: Roethlein, 1912; Shaw, 1969; Smither & Braun, 1994; Sheedy et al., 2005; Beier & Oderkerk, 2019a.

A study by Bernard et al. (2013) finds no advantage of bolder weights above the standard for normally sighted participants’ reading speed at the fovea and in the periphery (measured using RSVP). Decreases in reading speed are found at the largest stroke widths, with stronger detrimental effects in the periphery. The authors conclude that “contrary to the popular belief, reading speed does not benefit from bold text in the normal fovea and periphery” (Bernard et al., 2013, p.33). These results must be interpreted in the context of the test material employed within the experiment (Figure 12). Bernard et al. (2013) create a range of weights of Courier using freeware (FontForge). Bolder weights are “created as if we added extra layers of pixels around the letter-strokes of the standard Courier font”, and lighter weights “as if we removed layers of pixels” (Bernard et al., 2013, p.35). This results in the bolder letterforms having decreased letter spacing and compromised counter sizes, particularly in the boldest test material. As some of the test material deviates from the conventions of typeface design, the application of this research to design practice is limited.

0.27x the quick brown fox jumped over the lazy dog
0.72x the quick brown fox jumped over the lazy dog
1x the quick brown fox jumped over the lazy dog
1.48x the quick brown fox jumped over the lazy dog
1.89x the quick brown fox jumped over the lazy dog
3.04x the quick brown fox jumped over the lazy dog

Figure 12: A range of Courier weights designed in FontForge, tested by Bernard et al. (2013). Used with permission of Elsevier Science & Technology Journals, from *Vision Research*, The Effect of Letter-Stroke Boldness on Reading Speed in Central and Peripheral Vision, Bernard, J-B., Kumar, G., Junge, J. and Chung, S., 84, pp.33-42, Copyright (2013); permission conveyed through Copyright Clearance Center, Inc.

2.4.3.4 LETTER WIDTH

The relationship between stroke width and counter size is mediated by letter width. Specifically, increases in stroke width are accompanied by increases in letter width in order to maintain letter counter (section 2.3.3). Therefore, while this PhD research is focused on stroke width, letter width is also examined.

Optimum letter width values are not well understood, and inclusive typography guidelines often do not make recommendations regarding this characteristic. An exception is Lighthouse International's guidelines which discourage the use of condensed typefaces (Arditi, 2018). A *condensed* typeface has characters with a decreased letter width.

The relationship between stroke width, letter width, and legibility is exemplified by Roethlein's (1912) study. Using the distance method with normally sighted participants, Roethlein (1912) finds that Cheltenham Old Style Bold is more legible than the regular weight, however Cheltenham Condensed Bold is not (Figure 13). Roethlein's (1912) study is analysed within Chapter 4's scientific review (section 4.3.2).

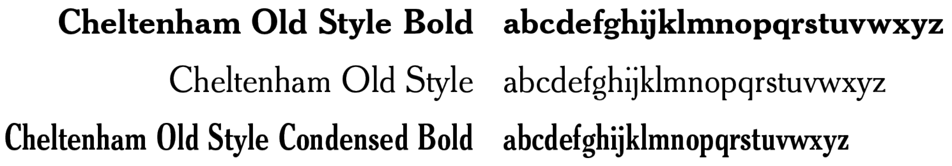


Figure 13: Cheltenham Old Style Bold, Cheltenham Old Style, and Cheltenham Old Style Condensed Bold, tested by Roethlein (1912).

Letter width has another important influence on legibility. As discussed in section 2.4.2, character size is a key typographic variable for low vision reading and a central aspect of inclusive typography guidelines. Evidence suggests that a horizontal increase in size (i.e. increased letter width) has a similar legibility benefit (Arditi, 2004).

2.4.3.5 MONOSPACED AND PROPORTIONALLY SPACED TYPEFACES

Inclusive typography guidelines recommend the use of monospaced versus proportionally spaced typefaces. *Monospaced* typefaces (e.g. Courier) have letter widths which are equal across all characters (Felici, 2012). In contrast, the characters within *proportionally spaced* typefaces (e.g. Times Roman) have unique widths. Continuous text is usually set in proportionally spaced typefaces.

RNIB and ISTD (2007) and Lighthouse International (Arditi, 2018) refer to the potential benefit of monospaced typefaces. CNIB’s (2020) guidelines specifically recommend choosing a monospaced typeface, however the typefaces they recommend—Arial and Verdana—are proportionally spaced, and a proportionally spaced typeface is used to typeset the guidelines. Confusion regarding this recommendation is also referred to by Tarita-Nistor et al. (2013). While monospaced typefaces differ in letter width compared to proportionally spaced typefaces, they are recommended in the context of letter spacing (e.g. CNIB, 2020; Arditi, 2018).

There is evidence to suggest that monospaced typefaces improve reading performance for people with low vision. Mansfield et al. (1996) find a reading acuity, reading speed, and CPS advantage for Courier Bold versus Times Roman for subjects with low vision (Figure 14). The study by Tarita-Nistor et al. (2013) finds an advantage for Courier near the acuity limit for participants with AMD (referenced in section 2.4.3.2). As these studies test commercial typefaces which differ simultaneously across characteristics, it is unclear what to attribute the legibility benefit of monospaced typefaces to. Mansfield et al. (1996, p.1493) acknowledge this stating that “Any of the differences between the fonts might be expected to influence reading performance.” The study by Mansfield et al. (1996) and the differences between Courier Bold and Times Roman are analysed in Chapter 4’s scientific review (section 4.3.6).

Courier Bold abcdefghijklmnopqrstuvwxyz
Times Roman abcdefghijklmnopqrstuvwxyz

Figure 14: Courier Bold and Times Roman, tested by Mansfield et al. (1996).

2.4.3.6 LETTER SPACING

Inclusive typography guidelines address the subject of letter spacing, however recommendations are quite general. Lighthouse International discourages “close letter spacing” (Arditi, 2018) and CNIB (2020, p.16) recommends keeping a “wide space between letters”. RNIB and ISTD (2007, p.52) advise against setting letter spacing “too tightly” or “too loosely”.

Evidence suggests that there is an advantage in reading performance for people with low vision if there is adequate spacing between letters (reviewed in Russell-Minda et al., 2007). This counteracts what is referred to as *crowding*; the interfering effect of adjacent letters (Legge, 2007). The crowding effect is more pronounced in peripheral vision (Bouma, 1970), which is of relevance to people with central vision loss who use peripheral vision for reading. Chung (2012) measures reading speed as a function of letter spacing for participants with central vision loss and finds the optimal letter spacing to be the standard spacing of the typeface (Figure 15).

Increased letter spacing beyond the standard is not found to improve reading performance. However, it should be noted that Chung (2012) varies the letter spacing of Courier, which is a monospaced typeface as discussed in section 2.4.3.5. As such, these results have limited application to design practice, as continuous text is usually set in proportionally spaced typefaces. In contrast, Beveratou (2016) finds a legibility benefit with increased letter spacing and leading of a sans serif proportionally spaced typeface, tested with partially sighted participants. *Leading* refers to the space between lines of text. Note that this benefit was not found in a proportionally spaced serif typeface. It was not reported whether this benefit was statistically significant.

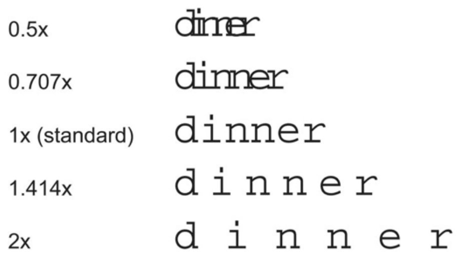


Figure 15: Five letter spacings tested by Chung (2012). Reprinted from *Optometry and Vision Science*, 89(9), Chung, S., Dependence of Reading Speed on Letter Spacing in Central Vision Loss, pp.1288-1298, Copyright © 2012 American Academy of Optometry, with permission from Wolters Kluwer Health, Inc. doi: 10.1097/OPX.0b013e318264c9dd

2.5 PRACTICE REVIEW: TYPEFACE DESIGN FOR LOW VISION

2.5.1 INTRODUCTION

While optimum typeface characteristics are not well understood, several typefaces are designed specifically for people with low vision. The typefaces APHont^{TM}, Tiresias^{TM} LPfont, Eido, Maxular Rx, and Matilda are introduced in the following sections.

2.5.2 APHONT AND TIRESIAS LPFONT

APHont^{TM} is developed by the American Printing House for the Blind (APH) for people with low vision. Features of APHont^{TM} include no serifs, heavier letters, wider letters, rounder letters, and even spacing between letters (Kitchel, 2004). Tiresias is a trademark of the RNIB font range and is designed by Laker Sharville (Figure 16). Tiresias^{TM} LPfont (i.e. Large Print Font) is “designed to have characters that are easy to distinguish from each other” (MyFonts, 2022). The typeface has small bracketed serifs and a larger stroke width compared to regular weight typefaces that are usually used to set continuous text.

Figure 16: APHont^{TM} developed by the American Printing House for the Blind (APH) and Tiresias^{TM} LPfont designed by Laker Sharville.

To my knowledge, there are no published experimental studies which test the reading performance of APHont^{TM}. The study by Rubin et al. (2006) tests Tiresias^{TM} LPfont and does not find a reading speed advantage compared to Times New Roman, Helvetica, and Foundry Form Sans. Conversely, Beveratou (2016) finds a reading speed advantage for Tiresias^{TM} LPfont compared to Times New Roman, Arial, Minuscule 6, Freight Micro Book, Optima, Century Gothic, and Palatino. It is not reported whether this difference is statistically significant. Beveratou (2016) suggests that the success of Tiresias^{TM} LPfont may be due to the typeface's stroke width. The study by Beveratou (2016) is included in Chapter 4's scientific review (section 4.3.8).

2.5.3 EIDO AND MAXULAR RX

Eido is designed by Jean-Baptiste Bernard, specifically for peripheral vision (Bernard et al., 2016) (Figure 17). The monospaced typeface is created "to reduce inter-letter similarity and consequently to increase peripheral letter recognition performance" (Bernard et al., 2016, p.1). Bernard et al. (2016) test Eido and Courier with normally sighted subjects, and find that Eido decreases perceptual errors in peripheral letter and word recognition. However, the study finds no significant difference in reading speed between the typefaces.

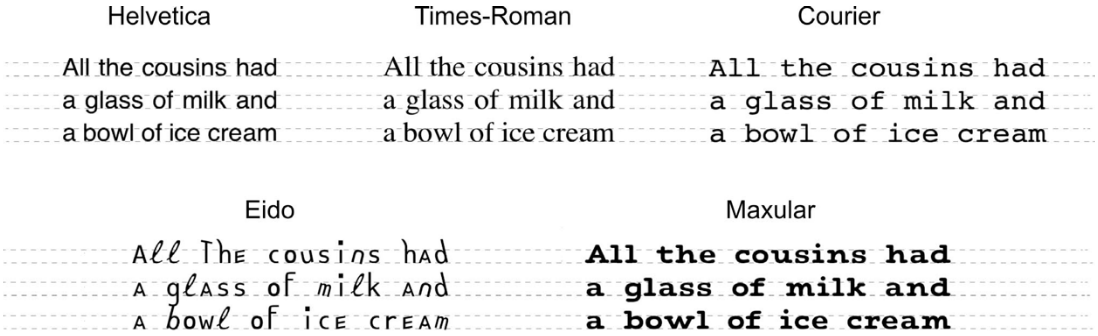


Figure 17: Helvetica, Times Roman, Courier, Eido, and Maxular Rx Bold, tested by Xiong et al. (2018). Attribution: "Figure 1. Demonstration of the five fonts used in the current study", Xiong, Y., Lorsung, E.A., Mansfield, J.S., Bigelow, C. and Legge, G.E., https://doi.org/10.1167/iovs.18-24334, CC BY-NC-ND 4.0.

Maxular Rx is designed by Steven Skaggs for people with macular degeneration (Delve Fonts, 2021) (Figure 17). This slab serif typeface has a large x-height and increased letter and word spacing (Delve Fonts, 2021). Xiong et al. (2018) test Eido and Maxular Rx Bold against Times Roman, Helvetica, and Courier. Compared with Times Roman and Helvetica, Eido permits smaller reading acuity, and Maxular Rx permits smaller reading acuity and CPS. Xiong et al. (2018) find that neither Eido nor Maxular Rx Bold perform better than Courier. The study by Xiong et al. (2018) is included in Chapter 4’s scientific review (section 4.3.9).

2.5.4 MATILDA

Matilda is designed by Ann Bessemans (2012), specifically for children with low vision (Figure 18). This is a particularly notable project as Bessemans’ (2012) typeface is created through interdisciplinary design research. Starting with a serif (DTL Documenta) and a sans serif (Frutiger) typeface, Bessemans (2016a) develops a laboratory typeface by modifying several parameters in order to assess their effect on legibility. Her research is focused on the concepts of homogeneity and heterogeneity. Bessemans 2016a) describes sans serif typefaces as homogeneous within their letterforms and heterogeneous within their rhythm, and the opposite for serif typefaces. Experimental results find that the serif typeface is more legible for children with normal vision, with the differences between the typefaces less pronounced in low vision children. Results suggest that for low vision children, a more irregular (i.e. heterogonous) rhythm is beneficial for reading. Based on this research, Matilda is developed; a typeface family which includes a serif, an italic, and a bold.

Figure 18: Matilda Regular, Bold, and Italic, designed by Bessemans (2012). Reprinted from *Letterontwerp voor kinderen met een visuele functiebeperking*, Bessemans, A., PhD, Leiden University and Hasselt University, Copyright (2012). Courtesy of Ann Bessemans © 2012.

2.6 SUMMARY

This chapter reviews inclusive typography guidelines, typeface legibility research, and typefaces designed specifically for low vision readers. The literature review demonstrates that while scientific research illustrates the influence of typeface characteristics on low vision legibility, optimal parameters are not well understood. This is reflected in the inclusive typography guidelines reviewed, which offer general recommendations. The lack of knowledge on optimal typeface parameters is also reflected in the varying features of typefaces designed specifically for low vision adults.

The literature review presents evidence that character size, presence or absence of serif, stroke width, letter width, monospaced versus proportional typefaces, and letter spacing influence legibility for low vision readers. While character size is known to be a critical typographic variable, text is generally not presented at sizes which are adequate for low vision reading. Typeface characteristics play an important role in legibility, particularly for people reading at their acuity limit.

Evidence suggests that sans serif typefaces are more legible for low vision readers (e.g. Morris et al., 2002). Stroke width and letter width influence legibility for low vision readers (e.g. Arditi, 2004), however optimum values are not well understood. Evidence suggests that monospaced typefaces improve reading performance for people with low vision (e.g. Mansfield et al., 1996), however it is not clear what typeface characteristics to attribute this performance benefit to. Evidence suggests that increased letter spacing may improve legibility for proportionally spaced sans serif typefaces (Beveratou, 2016), however this does not appear to benefit seriffed (Beveratou, 2016) or monospaced (Chung, 2012) typefaces.

My PhD research responds to the gap in knowledge regarding optimal typeface parameters for low vision adults through a scientific review. Chapter 4 presents a consolidation of scientific knowledge on stroke width. Distinct from a literature review, this is a quantitative analysis of ten legibility research studies. Chapter 4 addresses the PhD research question: What is the optimal typeface stroke width for low vision adults? The analysis entails measuring and visualising the stroke width and letter width of typefaces tested within legibility experiments. This generates knowledge on the relationship between quantified typeface characteristics and reading performance. Eight studies introduced in Chapter 2 are analysed in Chapter 4, which includes analyses of Tiresias™ LPfont (section 4.3.8) and Maxular Rx Bold (section 4.3.9) designed for low vision readers. The scientific review informs inclusive typography guidelines which—as reviewed in this chapter—recommend a range of typeface weights from regular to bold.

This chapter’s literature review highlights consistent issues concerning the controlled investigation of typeface parameters. Regarding stroke width, the studies by Arditi et al. (1995a) and Bernard et al. (2013) illustrate the limited application of research when test material is unconventional in its construction. The test material employed by Arditi et al. (1995a) is the only published example of a laboratory typeface designed for the controlled investigation of stroke width. My PhD research responds to this gap in knowledge through the creation of a design-informed laboratory typeface capable of the controlled study of stroke width and letter width. The design of this typeface was the initial focus of my PhD research and is presented in Chapter 5. The typeface was designed to address the PhD research question experimentally, however it was not tested with research participants. Instead, the value of my typeface design practice was that it gave rise to the quantitative analysis of typefaces, which became the focus of the PhD research.

Chapter 2 examines the context for this PhD research, reviewing the relevant literature and practice in the area of low vision legibility. Next, Chapter 3 focuses on the interdisciplinary research methodology of the PhD investigation, providing a foundation for the practice Chapters 4-6 which follow.

CHAPTER 3. METHODS AND METHODOLOGY: INTEGRATING SCIENTIFIC AND DESIGN KNOWLEDGE

3.1 INTRODUCTION

This PhD develops interdisciplinary knowledge on typeface legibility for low vision adults, contributing to future legibility research and inclusive typography guidelines. As reviewed in Chapter 2, optimum values for typeface characteristics are not well understood. Further, due to challenges regarding the controlled investigation of typefaces, there are limitations in the application of scientific research to design practice. This chapter describes the methodology for interdisciplinary typographic knowledge construction employed within this PhD research.

This chapter begins by reviewing the theoretical basis for an interdisciplinary approach to typeface legibility. The role of design knowledge in scientific research and scientific knowledge in design practice is considered. A framework for integrating scientific and design knowledge—*scientific design*—is introduced (Cross, 2001). Knowledge construction within design practice is also examined, and the study of ‘designerly ways of knowing’ through design artefacts—*design phenomenology*—is discussed (Cross, 2007).

The final sections of this chapter describe *research through design* and the methods of practice employed within the PhD research. The theoretical foundation of practice-based research is introduced, alongside the specific ways in which practice informs my research. Three methods of practice are discussed in greater detail including parametric typeface design, typeface measurement, and information visualisation. This chapter on methodology and methods lays the foundation for the following Chapters 4-6 which present the practice-based PhD research.

3.2 REVIEW OF INTERDISCIPLINARY APPROACHES TO TYPEFACE LEGIBILITY

3.2.1 DESIGN KNOWLEDGE WITHIN SCIENTIFIC RESEARCH

The importance of embedding design knowledge within scientific legibility research lies at the foundation of this PhD investigation. The initial focus of the PhD was to create a design-informed typeface for the controlled study of stroke width (presented in Chapter 5). This interdisciplinary approach was motivated by the lack of controlled experiments and design-informed test material in legibility research (section 2.3). The following paragraphs contextualise this interdisciplinary approach; the precursors to the PhD research, contributions of the PhD research, and developments in the discipline coinciding with the PhD research spanning more than a decade.

Design academics long ago warned that other disciplines would take on questions without a design perspective if the design community did not pursue research that crossed into visual communications expertise (Strickler, 1999). This describes the field of legibility research, as is evidenced by the predominance of scientist-led studies and consistent issues regarding test material (section 2.3). Before this PhD research was initiated in 2007, convincing arguments had been made for interdisciplinary approaches to typographic knowledge construction (Dyson, 1999; Lund, 1999). I was influenced by this, as well as Bigelow’s laboratory typeface (Morris et al., 2002) (section 2.3.3) which embedded design knowledge into scientific research.

In the first year of my PhD research, I made the case for interdisciplinary research within a lecture at ATypI (Association Typographique Internationale) 08 St Petersburg entitled *The Role of Typeface Design Within the Scientific Study of Legibility* (von Ompteda, 2008). Within this talk I argued that “practice-based typeface design research has a critical role to play in the scientific study of legibility” (von Ompteda, 2008). During the second year of my PhD, I again made the case for interdisciplinary legibility research within a paper entitled *Innovation in Inclusive Typography: A Role for Design Research* (von Ompteda, 2009a) (Appendix 1), presented at Include 2009; an international conference on inclusive design organised and hosted by the Royal College of Art Helen Hamlyn Centre. This paper critiqued legibility research methodologies, specifically the paucity of controlled investigation (e.g. Mansfield et al., 1996) and testing typefaces which are unconventional in their construction (e.g. Arditi et al., 1995a) (von Ompteda, 2009a). I assessed Arditi et al.'s 1995a study (section 2.3.3) as being unable to inform design practice, and practice-based design research as an opportunity to undertake investigations underpinning our discipline. The paper introduced this new context for legibility research, involving design researchers developing typefaces as test material and running experiments supervised by scientists. The paper concluded that design research—acting as an intermediary between scientific research and design practice—had a potentially exciting contribution to make to this field.

In reviewing the literature over a decade since the PhD began, the need for interdisciplinary approaches to typeface legibility research remains crucial. While there are increasing interdisciplinary contributions to legibility research, including design-informed laboratory typefaces (e.g. Bessemans, 2016a; Dyson & Beier, 2016), arguments for interdisciplinary approaches persist. Bessemans (2016b) references legibility studies that are typographically incorrect and argues for the collaboration or sharing of expertise between typographic design and scientific research. Dyson (2013) argues for an interdisciplinary approach between psychology and design, and specifically for involving designers in the development of test material to improve the applicability of scientific research to design practice. Beier (2016) argues for interdisciplinary collaborations involving scientists providing testing methods and analyses, and designers determining research questions and developing test materials.

My PhD research began in 2007, however the interdisciplinary approach that it takes remains relevant. The laboratory typeface presented in Chapter 5 demonstrates that experimental test material can be developed for the controlled study of stroke width while reflecting design knowledge in its construction. A design-informed laboratory typeface for the controlled study of stroke width is yet to be published in the research literature.

3.2.2 SCIENTIFIC KNOWLEDGE WITHIN DESIGN PRACTICE

The importance of scientific knowledge informing design practice equally lies at the foundation of this PhD investigation. This PhD research generates interdisciplinary typographic knowledge on low vision legibility with the purpose of influencing design practice. Similar to the previous section (3.2.1), such interdisciplinary approaches are rare. The following paragraphs contextualise this issue, addressing relationships between science and design, and specifically between scientific legibility research and typographic design practice.

Relationships between design and science have a long history of concern within the field of design research (Cross, 2001). One established relationship—*scientific design*—(Cross, 2001) is appropriate for the integration of scientific knowledge into typography practice. *Scientific design* refers to “modern, industrialized design—as distinct from pre-industrial, craft-oriented design-based on scientific knowledge but utilizing a mix of both intuitive and nonintuitive design methods” (Cross, 2001, p.52). In this context, science is seen as informing design tasks which have, over time, become “too complex for intuitive methods” (Cross, 2001, p.52). Cross (2001) does not believe that *scientific design* is a controversial concept, and simply reflects the reality of modern design practices like materials science.

Scientific design does not reflect the field of typography. A longstanding historical disconnect exists between scientific and design approaches to this subject area (Lund, 1999; Beier, 2016). Designers have questioned the positivist belief in experiments as the only method of generating valid knowledge, over their own knowledge articulated as tacit craft experience, professional knowledge, and visual sensibility (reviewed in Lund, 1999).

There has been a general consensus of the lack of success of much of the legibility research in the 20th century, which is considered to be obvious and as such not useful for the design community (Dyson 1999; Lund, 1999). The lack of impact of legibility research on the design community has persisted from the 20th century to the present (Beier, 2016).

The context of inclusive typography now makes it crucial for typographic practice to shift to *scientific design.* Applying Cross' (2001) writing to this subject: the task of designing for low vision readers is too complex for intuitive methods alone. While designers have long deemed scientific research as valuable in the 'special case' of visually impaired readers (e.g. McLean, 1980), inclusive typography is no longer a special case. Low vision has increased in prevalence (Mariotti, 2012) and inclusive design approaches have become more pervasive (Clarkson et al., 2003).

While the integration of scientific knowledge into typographic practice is not common, there are encouraging developments. The creation of inclusive typography guidelines (reviewed in section 2.4) which involve designers is a step in the right direction (e.g. RNIB & ISTD, 2007). Another example is Dyson's (2013) research which aims "to take an interdisciplinary perspective which synthesizes knowledge and bridges a gap between research findings and the skilled judgement of designers" (Dyson, 2013, p.272). Legge and Bigelow's (2011) interdisciplinary research integrates scientific and design knowledge, showing that the distribution of print sizes in historical and contemporary publications falls within the range over which text can be read at maximum speed. Beier (2012) takes a historical, design, and scientific perspective on typeface legibility for the purpose of informing design practice. Publications such as this bode well for a potential future where typographic designers' work is informed by scientific research.

3.3 KNOWLEDGE CONSTRUCTION IN DESIGN

This PhD research consolidates scientific knowledge, generates design knowledge, and develops interdisciplinary knowledge. The integration of scientific and design knowledge can be described as interdisciplinary in nature (versus multidisciplinary), as the aim is to generate new common knowledge (Barnes & Melles, 2007). The following paragraphs consider the theoretical issues regarding knowledge construction in the field of design.

Knowledge plays a fundamentally different role within scientific research and design practice. While the aim of scholarly research is the knowledge of truths, a designer’s goal is the production of artefacts (Heylighen et al., 2009). Knowledge is a means to design and an end for research (Heylighen et al., 2009). Heylighen et al. (2009) articulate that while the act of designing is increasingly acknowledged to be or involve some kind of knowledge production, making this explicit is not part of the design process. However, when research inquiries stretch across disciplinary boundaries, the onus is on the design community to formalise its own knowledge (Carvalho & Dong, 2009).

Knowledge associated with the practice of design is often referred to as experiential or tacit, based on Polanyi’s concept of the unarticulated and informal knowledge of experienced craftsmen and designers (Polanyi, 1976). In keeping, communication design has a long history of teaching through an atelier method, with students learning through the completion of design projects under the guidance of a professional (Strickler, 1999; Storkerson, 2008). In such a context, there is an understanding of a correct design solution but a lack of an underlying clarity of “how design works” (Storkerson, 2008, p.3). Communication design educators have been known to “complain about the lack of knowledge content in their field” (Storkerson, 2008, p.3).

While communication design has traditionally emphasised intuition and creativity over empirical research, Bennett (2006) foresees a future where designers are engaged in a research process. She argues that through "interdisciplinary research approaches, graphic designers can both question and affirm their intuitive inclinations" (Bennett, 2006, p.14). Walker (2017) argues that communication design research is thriving. She notes an established research tradition of determining the optimal effectiveness of design artefacts for different user groups. Walker (2017) also highlights strong practice-based research outputs within the discipline, including in the area of typeface design.

One of the major contributions of this PhD research is the formalisation of design knowledge. This is accomplished through the quantitative analysis of typefaces. Typefaces provide rich objects of study to investigate what Cross (2007) refers to as 'designerly ways of knowing'. As legibility is a goal when designers create a typeface for continuous reading (Dyson, 2013), my analysis of text typefaces formalises design knowledge on legibility. According to Cross' (2007) taxonomy of design research, a focus on design artefacts as opposed to designers or their processes is categorized as a design phenomenology study. *Design phenomenology* is defined as the study of the form and configuration of artefacts (Cross, 2007). According to Cross (2007, p.125) "we must not forget that design knowledge resides in *products* themselves: in the forms and materials and finishes which embody design attributes." He continues, "… we would be foolish to disregard or overlook this informal product knowledge simply because it has not been made explicit yet – that is a task for design research. So too is the development of more formal knowledge of shape and configuration – theoretical studies of design morphology."

3.4 PHD RESEARCH METHODS

3.4.1 INTRODUCTION

The following sections describe the methods employed within the PhD research. Research through design is introduced first, as this approach underlies the investigation. This is followed by an introduction to the methods of practice, discussed in the context of practice-based research. Lastly, the methods of the PhD research are described: parametric typeface design, the measurement of typefaces, and information visualisation. Statistical analyses are not described in this chapter, as they play a supporting role in the understanding of the visualisations in Chapter 6. Statistical analyses are described in section 6.3.1.2.

3.4.2 RESEARCH THROUGH DESIGN

The fundamental approach of this PhD investigation is *research through design* (RtD). This term has its roots in an oft-cited (Friedman, 2008) paper by Christopher Frayling published in the *Royal College of Art Research Papers* in 1993. He proposed three categories of research in art and design, one of which was “Research through art and design” which encompassed “the degree by project” he was familiar with at the Royal College of Art, with MPhil and PhD degrees involving studio work and a research report (Frayling, 1993, p.5). RtD is a practice-based form of inquiry, which is “now widely adopted in humanities-based research cultures and beyond” (RTD2017, 2017).

Practice-based research is defined as “an original investigation undertaken in order to gain new knowledge partly by means of practice and the outcomes of that practice” (Candy, 2006, p.1). For practice-based researchers, the creation of an artefact is central, with the making process providing “opportunities for exploration, reflection and evaluation” (Candy & Edmonds, 2018, p.66). This making process results in a transformation in the ideas, which then influence the creation of new artefacts (Candy & Edmonds, 2018).

The practice within my PhD research focuses first on typeface design and then on information visualisation. Through my typeface design practice, the methods for the quantitative analysis of typefaces emerge; the measurement and visualisation of typeface proportions. My practice knowledge in the areas of typography and typeface design inform the myriad of judgements throughout the research process. My knowledge of typography informs which typefaces to measure for the generation of design knowledge (section 6.2). My knowledge of typeface design informs which letters of a typeface to measure and how to measure them (section 4.2.2).

An important framework within practice-based research is *reflective practice*, which “attempts to unite research and practice, thought and action into a framework for inquiry which involves practice, and which acknowledges the particular and special knowledge of the practitioner” (Gray & Malins, 2004, p.22). This concept originates from Donald Schön’s *The Reflective Practitioner* (1983) and encourages reflection in different ways (Gray & Malins, 2004). *Reflection-in-action* “involves thinking about what we are doing and reshaping action while we are doing it” (Gray & Malins, 2004, p.22). An example of reflection-in-action within my PhD research is measuring typefaces tested in scientific experiments; an action which emerges intuitively without planning (section 1.6). In contrast, *reflection-on-action* is “part of the generic research process of review, evaluation and analysis” (Gray & Malins, 2004, p 22). Examples of reflection-on-action within my PhD include reflecting on the laboratory typeface in 2009 and changing the direction of the research to focus on information visualisation (section 1.6). The most important example of reflection on action is perhaps what lies before the reader in this moment – the writing of the PhD thesis itself.

3.4.3 METHODS OF PRACTICE: INTRODUCTION

Characteristic of art and design research is the use of multiple methods which are customised to meet the needs of an individual project (Gray & Malins, 2004). These methods are usually visual, and mainly derived from practice or adapted from other disciplines (Gray & Malins, 2004). The methods employed within my PhD research are typeface design, the measurement of typefaces, information visualisation, and statistical analyses. These represent design methods (typeface design), scientific methods (statistical analyses), and interdisciplinary methods (measurement of typefaces; information visualisation). The central method of my PhD is information visualisation, with the other methods playing a supporting role through the research process. It is noteworthy that “visualization – drawing (in all forms), diagrams” (Gray & Malins, 2004, p.30) is identified as a method of practice that can be effectively employed as a “robust and rigorous” method for “accessible and disciplined” inquiry by Gray and Malins (2004, p.29).

Information visualisation as a method of practice circumvents a contentious issue with regard to practice-based PhD research. While an artefact may represent new knowledge, this understanding must be communicated “in a form that meets the requirements of shared knowledge” (Candy & Edmonds, 2018, p.67). Ambiguity is central to this issue, as within the creative arts an artefact (e.g. painting) may be interpreted in different ways which is “fundamental to the nature of art” (Candy & Edmonds, 2018, p.67). The artefact therefore requires another (linguistic) method of communication, in order to frame the way it is viewed and the knowledge is understood (Candy & Edmonds, 2018). In contrast, information visualisation is a communication medium unto itself and aims to transmit knowledge unambiguously. Anyone with the skills needed to read a graph should be able to understand the knowledge it represents. While the accompanying PhD thesis supports and contextualises this understanding, information visualisation directly communicates the knowledge generated by the PhD research.

3.4.4 METHODS OF PRACTICE: PARAMETRIC TYPEFACE DESIGN

Parametric typeface design is introduced in Chapter 2’s discussion of laboratory typefaces as experimental test material (section 2.3.3). To reiterate, the characteristics of parametric fonts can be mathematically altered in order to produce an infinite variety of fonts (Knuth, 1986). Adapted to a legibility context, parametric laboratory typefaces can be adjusted systematically to allow for controlled investigation into the influence of specific parameters on reading performance. Such test material is important, because experiments testing commercial typefaces cannot assess the influence of isolated parameters on legibility. As discussed in Chapter 2, Arditi is well known for his research utilising parametric typefaces for controlled investigation of typeface characteristics (Arditi et al. 1995a; Arditi et al. 1995b; Arditi, 2004; Arditi & Cho, 2005) (section 2.3.3). However, because these laboratory typefaces are unconventional in their form, the results of these experiments have limited application to design practice. Chapter 2 also illustrates that a design-informed laboratory typeface can test the influence of an isolated parameter on legibility while remaining applicable to design practice. The study by Morris et al. (2002) is presented, employing test material created by typeface designer Bigelow which controls parameters other than serifs (section 2.3.3). As referenced in section 3.2.1, further design-informed laboratory typefaces have been created (e.g. Bessemans 2016a; Dyson & Beier, 2016).

At the time when this PhD began in 2007, a design-informed parametric typeface for the controlled investigation of stroke width on legibility had not yet been created. The parametric typeface I designed was inspired by Arditi's controlled experiments investigating typeface proportions (e.g. Arditi et al., 1995a) and Bigelow's design-informed test material (Morris et al., 2002) (section 2.3.3). I presented specific plans for the parametric typeface in my lecture at ATypI 08 St Petersburg (von Ompteda, 2008), and the completed typeface at the Interdisciplinary Graduate Conference 2009, held at the University of Cambridge (von Ompteda, 2009b). More recently, Beier (2018) designed a laboratory typeface for the controlled study of stroke width and letter width with inconclusive and unpublished results.

Chapter 5 presents my laboratory typeface, designed for the controlled investigation of stroke width. The typeface is created in FontLab using multiple master technology (section 5.2.4). As bolder typefaces generally have an increased letter width (Figure 3), the laboratory typeface allows for the controlled study of both parameters in order to reflect design practice.

The typeface is designed to test stroke width values found to have higher legibility (scientific knowledge) while reflecting text typeface proportions (design knowledge). Decisions regarding the parametric typeface's proportions are based upon the quantitative analysis of typefaces, involving the measurement of typefaces and information visualisation. As introduced in Chapter 1 (section 1.6), the typeface itself is less important than the methods for the quantitative analysis of typefaces which emerge from this design process. The measurement of typefaces and information visualisation are described in the following sections.

3.4.5 METHODS OF PRACTICE: MEASUREMENT OF TYPEFACES

As introduced in Chapter 1 (section 1.7), characterising typefaces entirely in quantitative terms is not common within design practice. Typeface weight names (e.g. bold), letter width names (e.g. condensed), and sizes (e.g. 12 point) do not translate into standardised typeface measurements. For example, two condensed bold typefaces set at 12 point will have different stroke widths, letter widths, and vertical heights. While typefaces are generally not characterised quantitatively, there are many examples of designers measuring typefaces. A 'neutral' typeface is designed by Kai Bernau (2005) based on detailed measurements of ten well known sans serif typefaces. Karen Cheng's (2005) book on typeface design includes many examples of measuring typeface parameters.

Within the scientific community, characterising typefaces entirely in quantitative terms is also not common. The work of Arditi (section 2.3.3) is an exception, which regularly characterises laboratory typefaces according to their parameters (e.g. stroke width of 30% cap height) (Arditi et al., 1995a). As introduced in Chapter 1 (section 1.7), scientists testing commercial typefaces regularly refer to test material using typeface names (e.g. Courier Bold) versus describing them quantitatively according to their parameters (e.g. Mansfield et al., 1996). Legge (2007, 112) describes a typeface as "a creative work, typically not amenable to quantitative stimulus description".

More recently, there are examples of legibility researchers describing typefaces quantitatively. For example, Xiong et al. (2018) measure the spacing of the typefaces they test, finding that spacing is a significant predictor of reading performance. This study is introduced in section 2.5.3 and described in more detail in section 4.3.9. While examples of quantifying typeface proportions exist in low vision legibility research, this approach remains uncommon.

This PhD research is focused on measuring the stroke width and letter width of typefaces, and to a lesser degree, letter spacing. Adobe Illustrator is used to measure each typeface's parameters. Typefaces are first outlined, and then measurements are taken by selecting the outline and using Illustrator's Transform tool to retrieve the height and/or width of the parameter, measured to a thousandth of a millimetre. In order to collect data on each parameter, specific typeface characters are chosen as representative. The rationale behind these more specific choices and further details regarding measurement techniques are described in Chapter 4 (section 4.2.2).

In order to consolidate scientific knowledge, the proportions of typefaces tested in legibility experiments are measured. In order to formalise design knowledge, the proportions of text typefaces utilised in design practice are measured. These measurements constitute typeface data. Note that within the context of information visualisation, typeface data (e.g. stroke width) is also referred to as a *variable*. The data is used to build the laboratory typeface based on scientific and design knowledge (section 5.3.3). The data is also analysed through information visualisation, described in more detail in the following section.

3.4.6 METHODS OF PRACTICE: INFORMATION VISUALISATION

3.4.6.1 EMERGENCE AND PRACTICE

Information visualisation was not initially planned as a research method, and instead emerged through the process of designing the laboratory typeface. As introduced in Chapter 1 (section 1.6), during the first year of PhD research I intuitively began measuring typefaces, created a spreadsheet of typeface data, and started graphing the data. This was done to gather and analyse data in order to build the parametric typeface based on scientific and design knowledge. During the second year of the PhD, it became clear that information visualisation was the central method of the research.

While collecting and graphing data came naturally to me based on my background in biology, I was also influenced by developments in the design world at the time. At the beginning of the PhD in 2007, information visualisation as a creative practice was exploding, fuelled by unprecedented access to information via the Internet, and by easier access to visualisation software (von Ompteda, 2019a). The Berlin-based publisher Gestalten published *Data Flow* in 2008; a survey of graphic design engagement in information visualisation (Klanten et al., 2008). Concurrent to the PhD research, I have been engaged in the information visualisation community through practice, teaching, and research (e.g. von Ompteda, 2019a; von Ompteda, 2019b).

Through the PhD research process, information visualisation developed from being an analytical tool to also becoming a communication tool. Data was in the zeitgeist, contributing to my belief that information visualisation would facilitate the dissemination of my research to the practicing design community. My goal was not only to communicate information, but to create a new lens through which designers might view the practice of typeface design. Information visualisation ultimately served an effective dissemination vehicle for my PhD research within the practicing design community, including a presentation I gave at *TYPO London* (von Ompteda, 2011b), an article I wrote for *Grafik* magazine (von Ompteda, 2011a), and the exhibition of my work at the *BIO.23 Biennial of Design* (2012).

While the PhD research employs information visualisation as a method, my practice is not approached from the information visualisation discipline. My visualisation practice is informed by my training in graphic design and biology, the latter routinely requiring the graphical representation of data. My graphic design training results in information visualisation work that differs in important ways from what I would have created based on scientific training alone. Graphic design training involves the development of a design sensibility, which is reflected in areas including typography, hierarchy, composition, and use of colour. None of the visualisations created within the PhD resemble those which are output by spreadsheet (e.g. Excel) or statistical (e.g. JMP) programs. These differences are described within the practice Chapters 4-6. While I did not approach this work from an information visualisation background, the PhD outcomes align with best practice within the field, for example the use of hierarchy (Yau, 2013).

3.4.6.2 CONTRIBUTION TO KNOWLEDGE

As introduced in section 3.4.6.1, information visualisation serves as an analytical and communication tool within the PhD research. As an analytical tool, information visualisation reveals patterns, relationships, and ultimately areas for further research (Hand, 2008). Visualising data is also useful for communicating findings to other people (Hand, 2008).

Typeface data is visualised through scatterplots, with stroke width data plotted on the *X* axis and letter width on the *Y* axis. These graphs analyse and communicate:

1) Scientific knowledge: Stroke width and letter width values of typefaces experimentally found to have higher and lower legibility;
2) Design knowledge: Stroke width and letter width values of sans serif text typefaces used in design practice;

3) Interdisciplinary knowledge: Relationship between stroke width and letter width values of typefaces found to have higher legibility (scientific knowledge) and typefaces utilised in design practice (design knowledge).

These scatterplots represent the main practice-based outcomes of the PhD research and contributions to knowledge. The visualisations of scientific knowledge address the research question: What is the optimal typeface stroke width for low vision adults? The visualisations of design knowledge constitute a design phenomenology study and serve to formalise design knowledge on typeface legibility. This analysis further clarifies the relationship between typeface weight names (e.g. bold) and numerical stroke width values. Visualisations of interdisciplinary knowledge elucidate how the proportions of typefaces found to have higher legibility (scientific knowledge) relate to the proportions of typefaces used in design practice (design knowledge). Interdisciplinary visualisations address the research question in the context of design practice, investigating optimal typeface weights for low vision adults.

3.5 SUMMARY

This chapter makes the case for an interdisciplinary approach to typeface legibility. The importance of integrating design knowledge into scientific research, and scientific knowledge into design practice is argued. Knowledge plays fundamentally different roles within science and design, and one of the major contributions of this PhD research is the formalisation of typeface design knowledge.

This PhD research is conducted *through* design, generating knowledge through practice itself and the outcomes of practice. Three primary methods of practice are employed in the research: parametric typeface design, typeface measurement, and information visualisation. My typeface design practice gives rise to the measurement and visualisation of typeface proportions; the quantitative analysis of typefaces. Visualisations of scientific, design, and interdisciplinary knowledge constitute the central practice-based outcomes of the research and contributions to knowledge.

This chapter presents the methodological foundation for the practice Chapters 4-6 which follow. The next chapter is focused on the consolidation of scientific knowledge on stroke width through the quantitative analysis of typefaces.

CHAPTER 4. SCIENTIFIC REVIEW: THE INFLUENCE OF STROKE WIDTH ON LEGIBILITY

4.1 INTRODUCTION

This chapter addresses the PhD research question: What is the optimal typeface stroke width for low vision adults? As reviewed in Chapter 2 (section 2.4.3.3), scientific evidence suggests that stroke width influences legibility, however optimal values are not well understood. As discussed in Chapter 3 (section 3.4.5), the typefaces tested in legibility studies are rarely described quantitatively, contributing to a lack of knowledge in this research area. This PhD investigation responds to this gap in knowledge through a quantitative analysis of typefaces.

This chapter presents a consolidation of scientific knowledge on stroke width in the context of low vision reading. Distinct from a literature review, this is a quantitative analysis of ten experimental studies which investigate the influence of stroke width on legibility. The studies include both scientific (e.g. Mansfield et al., 1996) and interdisciplinary design research (e.g. Beveratou, 2016). As in Chapter 2, the experimental studies apply to a high-resolution context and utilise both print and screen-based media for legibility testing. The testing methods are reported for each study.

The quantitative analysis involves measuring and visualising the stroke width and letter width of typefaces that are used as experimental test material in the ten studies. Analyses of letter spacing are also included in some cases. As discussed in Chapters 1-3, three typeface parameters—stroke width, letter width, and letter spacing—are examined because they vary alongside one another. The purpose of the analysis is to elucidate the relationship between typeface parameters and reading performance.

The following presents the methods employed in this scientific review, including the selection criteria for the ten studies, the forty typefaces analysed, and the methods of analysis (i.e. typeface measurement and information visualisation). The results of the investigation are presented for each study through visual (i.e. information visualisation) and written analyses of the typefaces used as experimental test material. The discussion includes analyses of stroke width and legibility, with recommendations for inclusive print guidelines and future research. Aspects of this chapter's scientific review informs the design of a laboratory typeface presented in Chapter 5.

4.2 METHODS

4.2.1 SELECTION OF SCIENTIFIC STUDIES AND TYPEFACES

Scientific knowledge is consolidated based on a review of experimental studies in which stroke width is a potential factor in the differing legibility of the typefaces tested. Some studies test stroke width directly, while others test a selection of typefaces and stroke width is considered by the authors to be a potential factor in their performance differences. All scientific studies are relevant to a low vision reading context, either testing typeface legibility with low vision participants, older participants, or normally sighted participants with low contrast or small sized typographic test material. Experiments using test material which does not resemble conventional typefaces (Arditi et al. 1995a; Arditi, 2004) (section 2.3.3), or does not add weight in the way that conventional typefaces do (Bernard et al., 2013) (section 2.4.3.3**)** are not included in the review. All studies either employ mixed-case or lowercase test material. One exception is the study by Sheedy et al. (2005) which tests uppercase in addition to lowercase.

Ten experimental studies meet the criteria outlined above. The final scientific studies included in the review (listed in chronological order) are: Roethlein (1912), Luckiesh and Moss (1940), Shaw (1969), Smither and Braun (1994), Mansfield et al. (1996), Sheedy et al. (2005), Beveratou (2016), Xiong et al. (2018), Beier and Oderkerk (2019a), and Beier and Oderkerk (2019b). This review of ten studies includes the analysis of forty typefaces (Table 1).

Table 1: Scientific studies included in the review and typefaces tested experimentally within each.

Scientific Study	**Typeface Tested**
Roethlein (1912)	Century Old Style
	Century Old Style Bold
	Cheltenham Old Style
	Cheltenham Old Style Bold
	Cheltenham Condensed Bold
Luckiesh & Moss (1940)	Memphis Light
	Memphis Medium
	Memphis Bold
Shaw (1969)	Gill Sans Roman
	Gill Sans Bold
	Plantin Roman
	Plantin Bold
Smither & Braun (1994)	Helvetica
	Helvetica Bold
Mansfield et al. (1996)	Times Roman
	Courier Bold
Sheedy et al. (2005)	Arial
	Arial Bold
	Verdana
	Verdana Bold
	Times New Roman
	Times New Roman Bold
	Georgia
	Georgia Bold
	Franklin Gothic Book
	Franklin Gothic Medium
	Franklin Gothic Demi
	Franklin Gothic Heavy

Table 1: Continued.

Scientific Study	**Typeface Tested**
Beveratou (2016)	Freight Sans Book
	Arial Regular
	Tiresias LPfont
Xiong et al. (2018)	Helvetica
	Times Roman
	Courier
	Maxular Rx Bold
Beier & Oderkerk (2019a)	Ovink Regular
	Ovink Semi Bold
	Ovink Ultra Black
Beier & Oderkerk (2019b)	Gill Sans Light
	KBH Text Regular

The choice of typefaces to measure is important, because typefaces with the same name from different foundries may not have the same proportions. Four scientific studies make reference to foundries (Roethlein, 1912; Luckiesh and Moss, 1940; Shaw, 1969; Mansfield et al.,1996), facilitating the choice of typefaces to measure. Once each typeface is located, it is superimposed over an image of the relevant study's published test material using Adobe Illustrator to ensure a match (Figure 19). Five scientific studies do not make reference to foundries (Smither & Braun, 1994; Sheedy et al., 2005; Beveratou, 2016; Xiong et al., 2018; Beier & Oderkerk, 2019b). In these cases, system fonts and commonly used foundries (e.g. Adobe) are explored, and each typeface is again superimposed over an image to ensure a match. Unfortunately, I am unable to find a match for Smither and Braun's (1994) versions of Century Schoolbook or Courier, and therefore only include Helvetica in the analysis. The study by Sheedy et al. (2005) is the only one which does not include an image of the test material, and therefore it is not possible to ensure a match between the typefaces I measure and the typefaces Sheedy et al. (2005) test. Beier and Oderkerk (2019a; 2019b) test a laboratory typeface (Ovink) and a proprietary typeface (KBH), therefore I measure these manually based on the studies' published test material. This is described in more detail in the following sections. All typefaces I measure are shown in the figures which accompany the analysis of each scientific study (section 4.3).

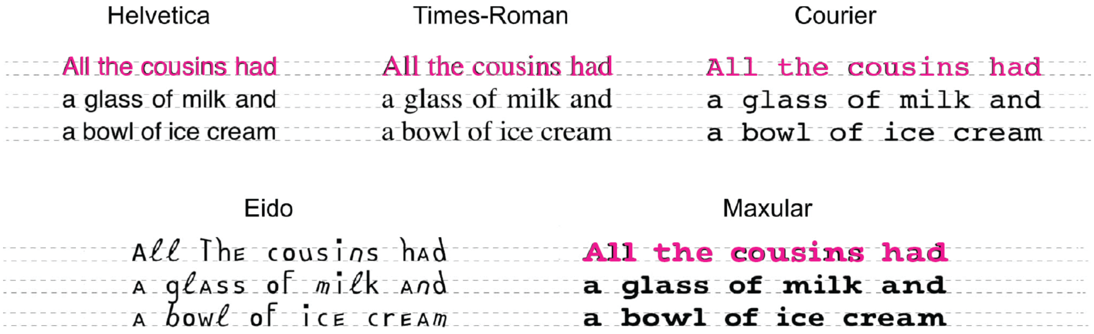


Figure 19: Typefaces I measure (pink) are superimposed over Xiong et al.'s (2018) test material. See Figure 17 for unmodified image. Attribution: "Figure 1. Demonstration of the five fonts used in the current study", Xiong, Y., Lorsung, E.A., Mansfield, J.S., Bigelow, C. and Legge, G.E., doi:https://doi.org/10.1167/iovs.18-24334, CC BY-NC-ND 4.0.

4.2.2 MEASUREMENT AND CALCULATION OF TYPEFACE PROPORTIONS

Data is gathered on forty typefaces' (Table 1) stroke width, letter width, and letter spacing. This entails first setting representative letters of each typeface at 10 point in Adobe Illustrator. These letters are then "outlined", essentially transforming each object from a typeface into a vector illustration. In some cases, letterforms are also modified in order to isolate the parameter of interest. For example, serifs are removed from outlines such that they do not unduly amplify measurements of stroke width and letter width, nor result in underestimates of letter spacing (see Figure 20). Outlines are then selected, and height or width information is retrieved using the "transform" tool in Adobe Illustrator. Measurements are recorded to a thousandth of a millimetre.

The following details regarding measurements are best understood with reference to Figure 1, which is reproduced from Chapter 1. As introduced in Chapter 1 (section 1.7), point sizes do not have an accurate numerical meaning. Therefore stroke width, letter width, and letter spacing data are calculated as a percentage of each typeface's x-height. The height of the lowercase 'x' is a common method to quantitatively describe the vertical height of a typeface (Legge, 2007).

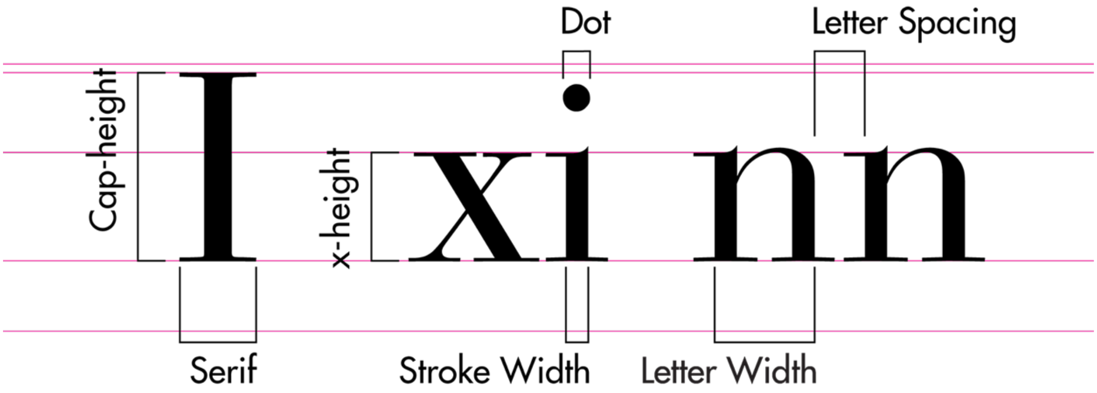


Figure 1: Typeface anatomy and letters employed for measurements. Figure reproduced from Chapter 1.

In order to gather data on the stroke width of each typeface, the lowercase 'i' is measured (Figure 20). This is a relatively simple letterform, consisting of a vertical stroke with a dot. The 'i' therefore requires little modification to prepare it for measurement. This typically involves deleting the dot, which is generally wider than the character's stroke width. Stroke width is calculated as a percentage of x-height.

Bigelow (2019) states that a more accurate estimate of a typeface's weight can be calculated through the more complex method of measuring ink area (or black pixel percentages). In contrast, my method focuses on stroke width, allowing for the efficiency required to measure the number of individual typefaces examined in the thesis, for example forty in Chapter 4 and 193 in Chapter 5. This method is aligned with Bigelow's (2019, p.166) statement:

"Type designers usually prefer to estimate weight easily by the ratio of vertical letter stem divided by x-height." My thesis reports this ratio as a percentage. For example, the stem/x-height ratio of 1:3.3 is equivalent to 0.30 (Bigelow, 2019), which is reported in my thesis as 30%.

Other legibility researchers also measure typeface weight as stroke width. For example, Arditi et al. (1995a) calculate stroke width as a percentage of cap height (section 2.3.3). Beier and Oderkerk (2019a) calculate stroke width as a ratio to ascender height (e.g. 'h') (section 4.3.10). My own calculation of stroke width is aligned with design practice as stated by Bigelow (2019), with Bernau's (2005) 'neutral' typeface serving as an example (i.e. stroke width calculated as a percentage of x-height).

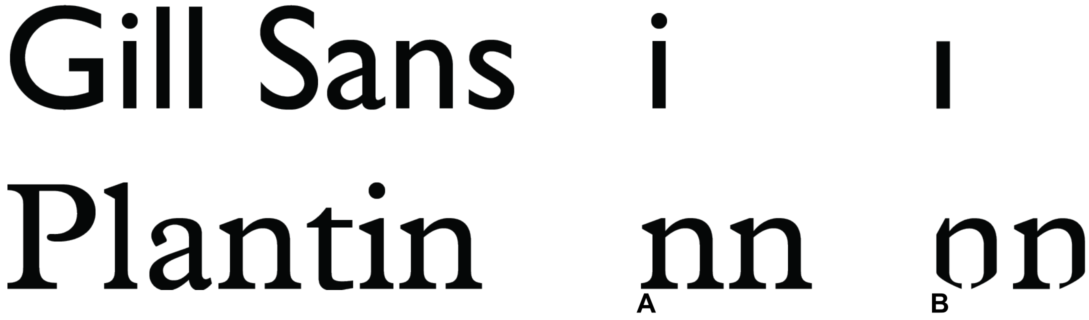


Figure 20: Gill Sans and Plantin characters modified for measurement. (A) Gill Sans 'i' and Plantin 'n's unmodified. (B) Gill Sans 'i' with dot removed for the measurement of stroke width; Plantin 'n' (left) with anchor points of serifs on both sides of the character removed for the measurement of letter width; Plantin 'n's with anchor points of serifs removed on the outer sides of both letters for the measurement of letter spacing. Note that when anchor points are removed from a vector outline, the character's appearance changes as can be seen with both 'n's.

In order to measure each typeface's letter width, the lowercase 'n' is chosen as representative (Figure 20). It is customary to use the lowercase 'n' and 'o' to determine the spacing of a typeface (Tracy, 1986; Cheng, 2005), therefore these characters represent the standard relationship between form and counterform. The 'n' is used because it has vertical strokes and is thus more comparable across typefaces of different styles. In contrast, the ‘o’ varies more in width and weight, depending on the shape of its curves. Bernau (2005) also utilises the letter ‘n’ in his measurement and analysis of typeface letter width proportions. Based on the same reasoning, letter spacing is measured as the distance between two lowercase ‘n’s (Figure 20). Letter spacing is calculated by measuring the entire width taken up by two ‘n’s (including the letter space between them), and then subtracting the letter width of two ‘n’s, resulting in a measurement of the letter spacing between the ‘n’s. Letter width and letter spacing are calculated as a percentage of x-height.

In order to analyse two of the studies, typeface proportions are measured based on published images of test material versus the typefaces themselves. This method is employed in the analyses of Beier and Oderkerk’s (2019a; 2019b) studies because they test typefaces that are not commercially available. Measuring these typefaces requires importing images of test material into Adobe Illustrator, drawing a rectangular shape within the area to be measured, and using the “transform” tool to retrieve the height or width as relevant (Figure 21). The x-height of Ovink Regular, Semi Bold, and Ultra Black (Beier & Oderkerk; 2019a) is measured using the height of the diagonal of the lowercase ‘k’. The stroke width and letter width of the typefaces are measured using the stem and the width of the ‘n’, respectively. The x-height of KBH Text (Beier & Oderkerk, 2019b) is measured using the stem of the lowercase ‘i’, and the stroke width and letter width are measured using the ‘i’ and ‘n’, respectively. Letter spacing is measured as the distance between the ‘u’ and the ‘m’. Because Beier and Oderkerk (2019b) also test a commercially available typeface (Gill Sans Light), it is possible to compare the accuracy of my manual measuring method. The accuracy is high, with stroke width identical (to three significant digits) and letter width differing by 0.1%.

Redacted

Figure 21: Example of my manual measurement of the typefaces tested by Beier and Oderkerk (2019a). The pink rectangles measure stroke width and letter width using the ‘n’, and x-height using the ‘k’. Ratios on the right side of the figure represent the stroke width to ascender height (e.g. ‘h’) ratio of each typeface. See section 4.3.10 for further information. See Figure 41 for unmodified image. This article was published in *Acta Psychologica*, 199, Beier, S. and Oderkerk, C.A.T., Smaller Visual Angles Show Greater Benefit of Letter Boldness Than Larger Visual Angles, pp.1-8, Copyright Elsevier (2019).

Once typeface measurement data is recorded, the relevant calculations are executed in Numbers; Apple’s spreadsheet application. As stated, typeface stroke width, letter width, and letter spacing are calculated as a percentage of x-height. In some cases, the difference in parameter values across typefaces is also calculated, by subtracting the parameter values of one typeface from another. While measurements are recorded to three significant figures (i.e. a thousandth of a millimetre), stroke widths and letter widths are recorded to two significant figures. *Significant figures* refer to “digits in a number that denote the accuracy of the measurement” (Zar, 1999, p.6). Therefore, while measurement accuracy allows for recording of stroke width data with three significant figures (e.g. 20.1% x-height), data is rounded to the nearest percent (e.g. 20% x-height). This level of accuracy is sufficient to achieve the goals of the research.

4.2.3 VISUALISING TYPEFACE DATA

As introduced in Chapter 3 (section 3.4.6.2), typeface data is visualised through scatterplots, with stroke width on the *X* axis and letter width on the *Y* axis (e.g. Figure 25). As information visualisation serves not only an analytical but also a communication function, these graphs differ from those that would be produced by spreadsheet (e.g. Excel) or statistics (e.g. JMP) programs in important ways. By plotting the data manually in Adobe illustrator, I am able to have control over all visual aspects of the graphs. This includes for example the weight of lines, typography, and use of colour. The most notable difference between my graphs and those that would be produced using applications like Excel, is the representation of the data itself. Instead of plotting a small ‘dot’ (the default), the data associated with each typeface is plotted using the lowercase ‘o’ of that typeface. Lowercase ‘o’s are chosen because they are similar to dots, relatively simple in form, and communicate much more meaning. As stated in section 4.2.2, ‘o’s are representative of the standard relationship between form and counterform within a typeface. While ‘n’s could be plotted, this would result in an overly complex visualisation, which would be more difficult for the viewer to understand. The final visualisations offer viewers an opportunity to examine actual letterform proportions associated with each data point, as the ‘o’ communicates the weight of the strokes, the width of the letter, and the consequent size of the counter.

All ‘o’s are plotted at the same point size, versus being scaled and presented at equal x-heights. Standardising x-height is not necessary to meet the purpose of the analysis, which is to elucidate the relationship between typeface parameters and reading performance. Therefore the positioning of each data point on the graph (i.e. the typeface’s stroke width and letter width values) is of central importance, while the relative size of the ‘o’s that are plotted is not. Beier and Oderkerk’s (2019a; 2019b) test material is visualised using commercial typefaces (see following paragraph), and this again has a negligible impact on the final visualisations which are focused on each typeface’s stroke width and letter width values.

In order to visualise the data from Beier and Oderkerk’s (2019a; 2019b) studies, each typeface’s data is plotted using ‘o’s from sans serif typefaces that have similar proportions. Beier and Oderkerk’s (2019a) typefaces are visualised as follows: Ovink Regular with Avenir Book (equivalent stroke width and 3% larger letter width); Ovink Semi Bold with Monotype Grotesque Bold (equivalent stroke width and 2% smaller letter width); Ovink Ultra Black with Gill Sans Ultra Bold (2% smaller stroke width and 8% larger letter width) (Figure 22). Beier and Oderkerk’s (2019b) KBH Regular is visualised with Avenir Regular (1% heavier stroke width and 5% smaller letter width) (Figure 23).

Redacted

A

abcdefghijklmnopqrstuvwxyz
abcdefghijklmnopqrstuvwxyz
abcdefghijklmnopqrstuvwxyz

B

Figure 22: Typefaces with similar proportions which are plotted in the place of Beier and Oderkerk’s (2019a) laboratory typeface. (A) Beier and Oderkerk’s (2019a) test material, with superimposed ‘n’s and ‘o’s from typefaces with similar proportions. Test material is discussed in more detail in section 4.3.10. This article was published in *Acta Psychologica*, 199, Beier, S. and Oderkerk, C.A.T., Smaller Visual Angles Show Greater Benefit of Letter Boldness Than Larger Visual Angles, pp.1-8, Copyright Elsevier (2019). (B) Typefaces with similar proportions: Avenir Book, Monotype Grotesque Bold, and Gill Sans Ultra Bold.

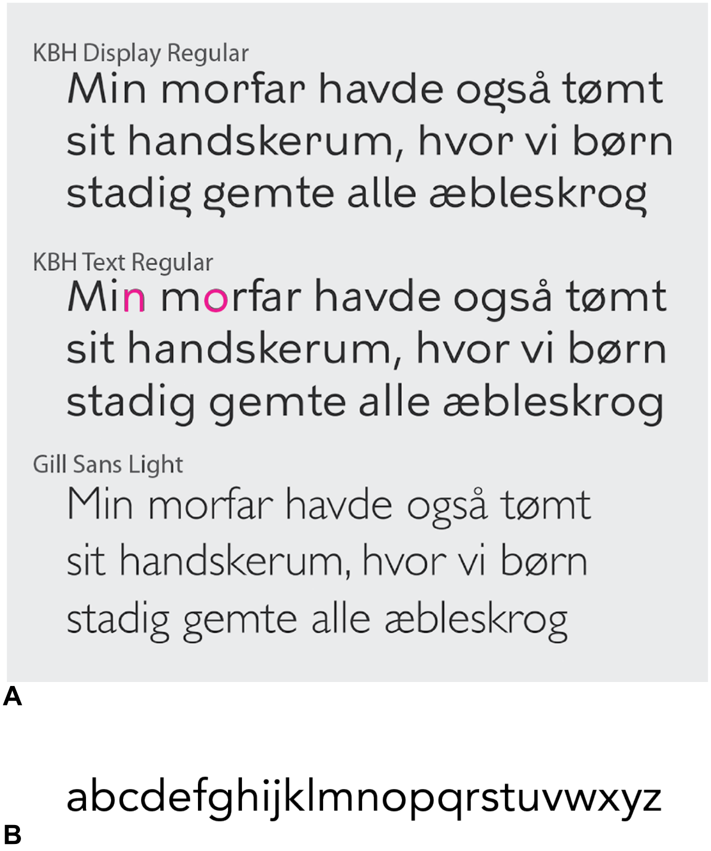


Figure 23: Typefaces with similar proportions which are plotted in the place of Beier and Oderkerk's (2019b) proprietary typeface KBH Text Regular. (A) Beier and Oderkerk's (2019b) test material, with superimposed 'n's and 'o's from Avenir Regular which has similar proportions. (B) Avenir Regular. Test material is discussed in more detail in section 4.3.11. Attribution: Figure 1 from the journal article: Beier, S. and Oderkerk, C.A.T., 2019. The Effect of Age and Font on Reading Ability. *Visible Language,* 53(3), pp.50-68, Copyright (2019), with permission from Visible Language.

Colour is also important in the design of the visualisations. CMYK colours are utilised due to their familiarity, vibrancy, and conceptual connection to a print context. CMYK is the colour profile utilised for print, and is an acronym representing the colours cyan (C), magenta (M), yellow (Y), and black (K).

Representing each experimental study, typefaces that are found to be more legible are presented as black (100%K), and those that are less legible are presented as grey (55%K) (e.g. Figure 25). This communicates a clear difference between typefaces, while also achieving adequate contrast with the background. This choice is also successful conceptually, as the less legible typefaces are literally less legible, with lower contrast between the grey ‘o’ and the white background.

The difference between more (black) and less (grey) legible typefaces is statistically significant only in studies which report statistics (e.g. Figure 30). A total of seven studies report the results of statistical analyses: Shaw (1969), Smither and Braun (1994), Mansfield et al. (1996), Sheedy et al. (2005), Xiong et al. (2018), Beier and Oderkerk (2019a), and Beier and Oderkerk (2019b). Three studies do not report statistics: Roethlein (1912), Luckiesh and Moss (1940), and Beveratou (2016).

Black lines are used to join the ‘o’s, connecting typefaces that are tested within each study. A line connecting a grey ‘o’ and a black ‘o’ represents that the former is less legible than the latter. In some cases this difference is statistically significant (e.g. Figure 30), while in other cases, statistics are not reported (e.g. Figure 25). In two visualisations, a line connecting two black ‘o’s is used to represent increasing legibility with increasing stroke width, with the difference between the two typefaces either not statistically significant (Figure 36), or relatively small if statistics are not reported (Figure 28). In one visualisation (Figure 38), a line connecting two grey ‘o’s is used to indicate that legibility increases with increasing stroke width, however both grey ‘o’s are less legible than the black ‘o’ in the visualisation (statistics not reported). Lastly, pink lines are utilised to create quadrants, which are described in the discussion (section 4.4.2) (e.g. Figure 47).

4.3 RESULTS

4.3.1 INTRODUCTION TO QUANTITATIVE ANALYSIS OF STUDIES

Each scientific study included in the review is described in the following sections. The accompanying figures show the typefaces tested in each study (e.g. Figure 24), and the analysis of their stroke width and letter width (e.g.

Figure 25). These two figures are designed to be viewed together, as the typefaces are presented in an order that reflects their position on the graph as much as possible. For example, the bolder typefaces are presented above the lighter typefaces (e.g. Figure 24), as they are generally visualised higher on the graph due to their larger letter widths (e.g. Figure 25). Typeface parameter values based on my measurements are also presented in accompanying tables (e.g. Table 2). The tables support an understanding of the visualisations, as they present the data that is graphed. The tables further include letter spacing data, which is not visualised.

While the visualisations are focused on stroke width and letter width, letter spacing is also discussed. Results from scientific studies are presented alongside my own calculations. My calculations are presented as percentages of x-height, for example: stroke width (22%). Whenever such percentages are reported, they refer to percentages of x-height, however the latter is not written. For example, a stroke width of 22% refers to a typeface that has a stroke width value that is 22% of its x-height. Comparisons between typefaces are also presented, for example: stroke width (5% lower). Stroke width, letter width, and letter spacing data for each typeface is also listed in Appendix 2.

When presenting each scientific study, this review does not report all experimental results. Assessments of legibility using maximum reading speed are not included in this review (e.g. Xiong et al., 2018). Maximum reading speed does not address size limitations of type and can produce different typeface legibility results than testing reading acuity or CPS (Xiong et al., 2018). Results are also not reported for aspects of studies testing normally sighted people at standard type sizes (Luckiesh & Moss, 1940; Beier & Oderkerk, 2019a), or test material lacking ecological validity (Beier & Oderkerk, 2019a). Typeface legibility assessments based on subjective evaluations (Smither & Braun, 1994) are also not reported. Blink rate data is not reported (Luckiesh & Moss, 1940), as this method is no longer used and its validity is described as questionable (Subbaram, 2004) with regard to its measure of legibility.

While some studies test typeface legibility with both low vision and normally sighted participants (e.g. Mansfield et al., 1996), only results associated with low vision participants are reported. One study tests both adults and children (Shaw, 1969), and only the results related to adults are reported here. Some studies also test older and younger adults (e.g. Smither & Braun, 1994), and only the results for older participants are reported.

In a number of studies, several factors influencing legibility are tested. Wherever possible, this review focuses on the testing of typeface stroke width, however in some cases there are statistical interactions with other factors, which are discussed when this arises. Lastly, while all studies use either mixed-case or lowercase test material, one study additionally uses uppercase (Sheedy et al., 2005), which is addressed in the relevant analysis (section 4.3.7).

The results of statistical analyses are presented throughout this review when available. If studies do not report statistical analyses, data from these studies is presented in Appendix 3. Appendix 3 reproduces data from the studies by Roethlein (1912), Luckiesh and Moss (1940), and Beveratou (2016).

4.3.2 ROETHLEIN, 1912

Roethlein (1912) tests twenty-six individual typefaces with six normally sighted participants. Legibility is assessed based on the distance method; this technique is similar to visual acuity testing (Subbaram 2004). The results for two typeface families—Century Old Style and Cheltenham Old Style—which elucidate the influence of stroke width on legibility are presented here (Figures 13 and 24). The study by Roethlein (1912) is introduced in Chapter 2 (section 2.4.3.4).

The legibility of typefaces is assessed based on the farthest distance from which isolated lowercase letters are identified. The twenty-six lowercase letters of each typeface are printed on a piece of paper at 10 point, presented individually (3.7cm between each letter), and arranged in random sequence. The paper is mounted on a sliding object that can present the letters at a variable distance from the subjects. The paper is held in place by a sheet of glass and is backlit. The procedure begins with the object at one end of a 440cm long bench and the subject at the other end. The subject attempts to identify each letter, after which the object is moved 20cm closer to the subject, and the subject once again attempts to identify the letters. This procedure is continued until every letter on the paper has been identified. The subjects are tested with each typeface twice, with a different arrangement of the letters. The farthest distance at which each letter is identified is recorded.

Distances are averaged across the twenty-six lowercase characters for each typeface. Data from Roethlein (1912) is reproduced in Appendix 3a. Typefaces with larger average distances are considered to be more legible, meaning that letters can be identified from further away. Based on the data, Century Old Style is less legible than Century Old Style Bold (Figures 24 and 25). Cheltenham Old Style is less legible than Cheltenham Old Style Bold (Figures 13 and 26). Further, Cheltenham Condensed Bold is less legible than Cheltenham Old Style Bold (Figures 13 and 26). The two typeface families are visualised separately (Figures 25 and 26) because they are discussed separately in Roethlein’s study. No statistical analyses are reported.

Based on the results, Roethlein writes that “legibility is very much increased by increased heaviness of face” (Roethlein, 1912, p.25). As presented in Chapter 2 (section 2.3.1), typefaces with increased stroke width generally have an increased letter width. While Century Old Style Bold has a heavier stroke width (10% higher), it also has an increased letter width (21% higher) compared to Century Old Style (Table 2; Figure 25). Similarly, Cheltenham Old Style Bold has a heavier stroke width (10% higher), as well as an increased letter width (18% higher) compared to Cheltenham Old Style (Table 2; Figure 26).

Cheltenham Old Style Bold is more legible than Cheltenham Old Style, however Cheltenham Condensed Bold is not (Appendix 3a). In reference to this result, Roethlein states that “whatever advantage might have been derived from increased heaviness of face, as compared with Cheltenham Oldstyle [sic], is neutralized by a disadvantage which is due to a narrowing of the internal spaces within the letters, and a consequent sacrifice of detail” (Roethlein, 1912, p.25). Based on my analysis (Table 2; Figure 26), Cheltenham Condensed Bold has a heavier stroke width (5% higher) than Cheltenham Old Style, and importantly, a decreased letter width (7% lower). Because the increased stroke width of Cheltenham Condensed Bold is not accompanied by letter width proportions that maintain an open counter, legibility is negatively impacted. Lastly, while letter spacing data is presented in Table 2, it is not discussed here because Roethlein (1912) presents letters individually, and therefore typeface letter spacing does not influence the results of the experiment.

Table 2. Scientific review parameter data: Roethlein, 1912. The table presents typefaces tested experimentally within the study, and their parameter values based on my measurements.

Typeface Family	**Typeface Tested**	**Parameter Values (% x-height)**		
		Stroke Width	**Letter Width**	**Letter Spacing**
Century Old Style	Century Old Style Bold	28	96	35
	Century Old Style	18	75	40
Cheltenham Old Style	Cheltenham Old Style Bold	31	100	31
	Cheltenham Old Style	21	82	35
	Cheltenham Old Style Condensed Bold	26	75	27

Century Old Style Bold abcdefghijklmnopqrstuvwxyz
Century Old Style abcdefghijklmnopqrstuvwxyz

Figure 24: Century Old Style Bold and Century Old Style, tested by Roethlein (1912).

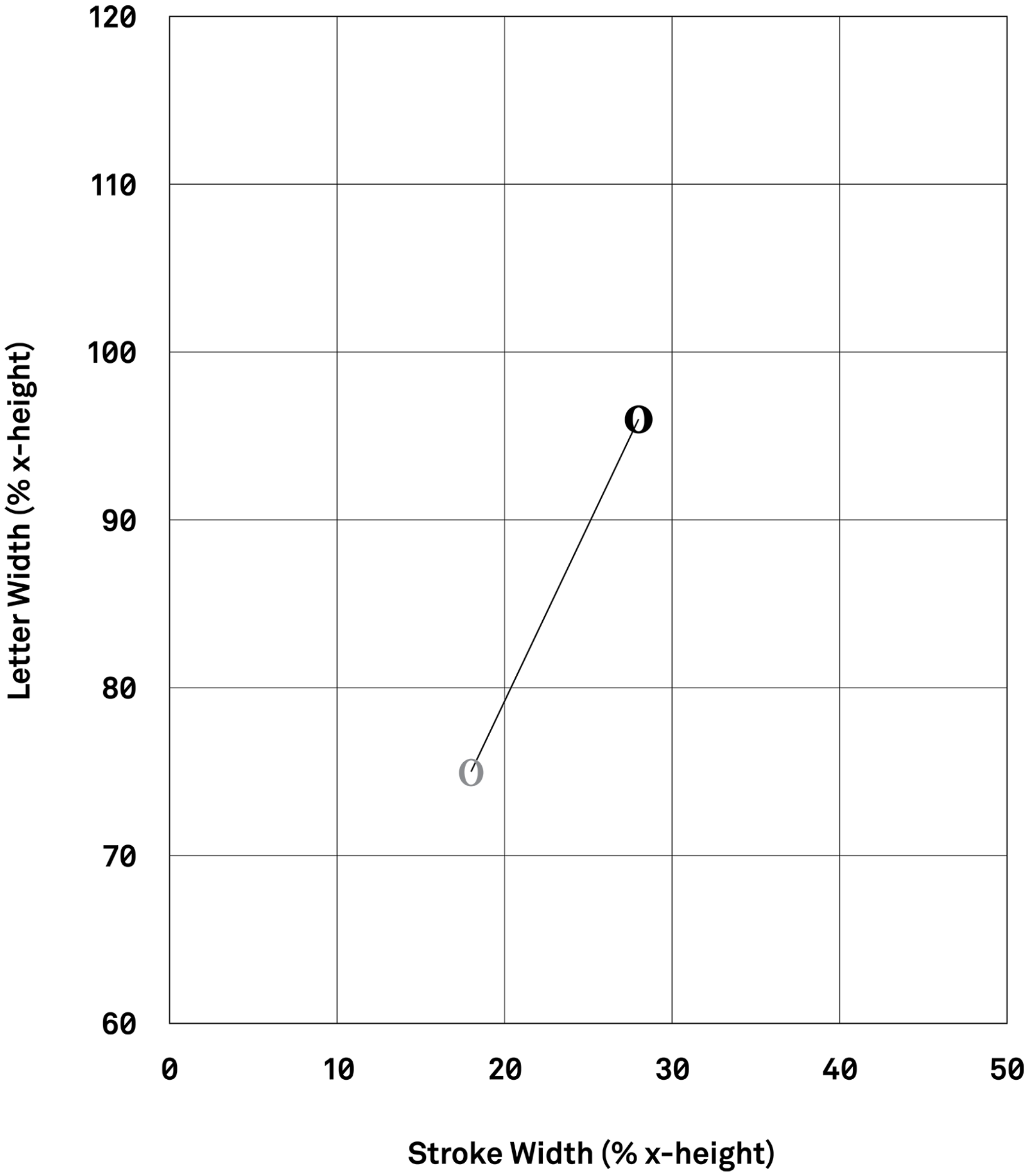


Figure 25: Century Old Style is found to be less legible (distance method) than Century Old Style Bold (Roethlein, 1912). Each typeface is represented by its lowercase ‘o’, with the colour grey representing lower legibility and black representing higher legibility. Typefaces found to be more and less legible are connected with a black line. The same graphic techniques are employed throughout the chapter.

Cheltenham Old Style Bold	**abcdefghijklmnopqrstuvwxyz**
Cheltenham Old Style	abcdefghijklmnopqrstuvwxyz
Cheltenham Old Style Condensed Bold	**abcdefghijklmnopqrstuvwxyz**

Figure 13: Cheltenham Old Style Bold, Cheltenham Old Style, and Cheltenham Old Style Condensed Bold, tested by Roethlein (1912). Figure reproduced from Chapter 2.

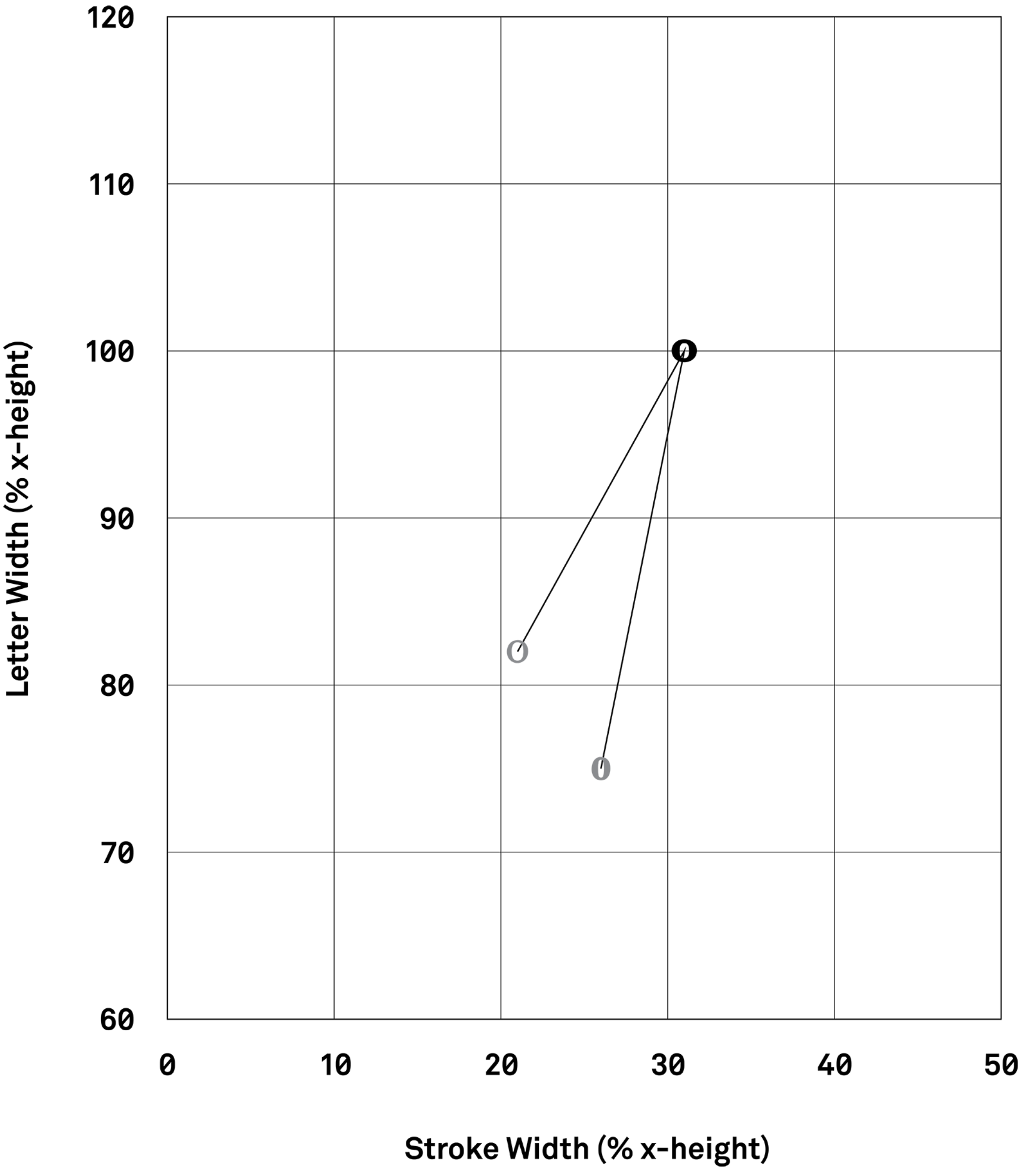


Figure 26: Cheltenham Old Style and Cheltenham Condensed Bold are found to be less legible (distance method) than Cheltenham Old Style Bold (Roethlein, 1912).

4.3.3 LUCKIESH AND MOSS, 1940

Luckiesh and Moss (1940) investigate the legibility of Memphis Bold, Medium, and Light (Figure 27). The typefaces are tested with five subjects using the Luckiesh-Moss visibility meter to determine legibility. Memphis Extra-Bold is also tested in the study, however it is not included in the analysis because this typeface has a notably larger x-height compared to the other weights in the typeface family, as the authors acknowledge.

The test material consists of continuous text (mixed case) set in Memphis Bold, Medium, and Light, printed at 10 point. Using the Luckiesh-Moss visibility meter, legibility is measured as visibility (Luckiesh & Moss, 1939). The meter is held by the subject in approximately the same position as eyeglasses. It consists of two gradient filters that are rotated simultaneously in front of the eyes while looking at the test material. The filters vary from almost clear to very dark, adjusting the light intensity of the stimulus and its background. The filters are rotated until the threshold of the visual task is reached, meaning that the test material can be recognised. Ten visibility measurements taken from five subjects (each tested twice), are averaged, and reported as a measure of relative visibility. Relative visibility data from Luckiesh and Moss (1940) is reproduced in Appendix 3b.

Luckiesh and Moss (1940) report that relative visibility increases with typeface weight. Increasing values are recorded for Memphis Light, Memphis Medium, and Memphis Bold, respectively (Appendix 3b). Regarding an optimum typeface weight, the authors state that “it appears that a practical optimum in boldness is obtained with Memphis Medium since the next step in boldness produces only a slight increase in relative visibility” (Luckiesh & Moss, 1940, p.175). They further state that “such an increment in visibility is not significant from statistical or practical viewpoints” (Luckiesh & Moss, 1940, p.176). No statistical analyses are reported.

Based on the study, Memphis Light is less legible than Memphis Medium and Memphis Bold when measured as visibility (Figures 27 and 28). As presented in Chapter 2 (section 2.3.1), typefaces with increased stroke width generally have an increased letter width and decreased letter spacing. My analysis of Memphis Bold and Memphis Light (Table 3; Figure 28) illustrates this typical relationship. Memphis Bold has a heavier stroke width (14% higher) as compared to Memphis Light, an increased letter width (6% higher), and decreased letter spacing (8% lower). These calculations suggest that visibility is improved by increased stroke width and/or letter width, despite a decrease in letter spacing.

My analysis of Memphis Medium and Memphis Light (Table 3; Figure 28) illustrates an atypical relationship between typeface stroke width and letter width. Memphis Medium has a heavier stroke width (8% higher) as compared to Memphis Light, decreased letter spacing (2% lower), and unusually, a decreased letter width (1% lower). These calculations suggest that the increased legibility of Memphis Medium is due to stroke width alone.

Table 3. Scientific review parameter data: Luckiesh & Moss, 1940. The table presents typefaces tested experimentally within the study, and their parameter values based on my measurements.

Typeface Tested	**Parameter Values (% x-height)**		
	Stroke Width	**Letter Width**	**Letter Spacing**
Memphis Bold	25	92	41
Memphis Medium	19	85	47
Memphis Light	11	86	49

Memphis Bold abcdefghijklmnopqrstuvwxyz
Memphis Medium abcdefghijklmnopqrstuvwxyz
Memphis Light abcdefghijklmnopqrstuvwxyz

Figure 27: Memphis Bold, Memphis Medium, and Memphis Light, tested by Luckiesh and Moss (1940).

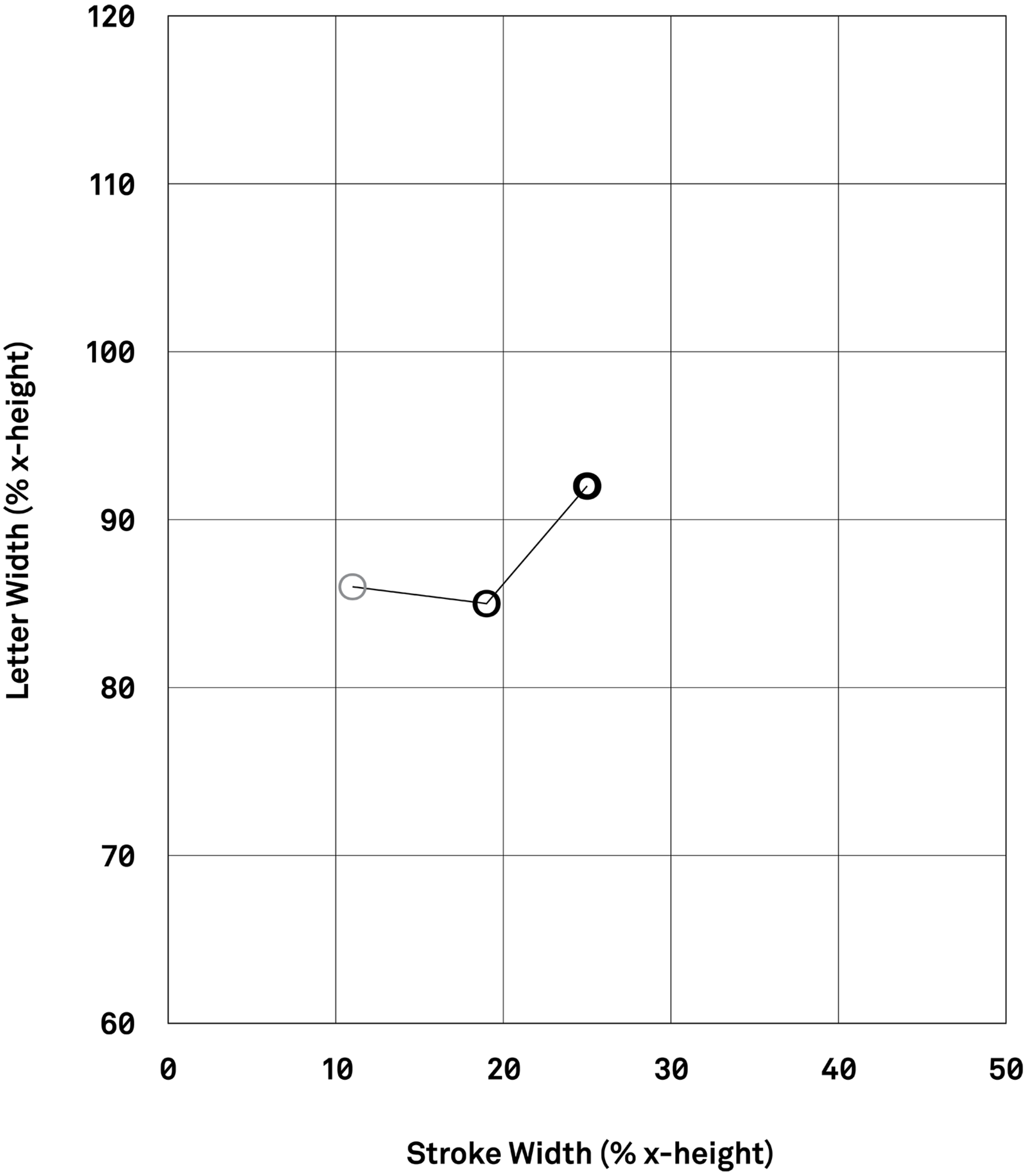


Figure 28: Memphis Light is found to be less legible (visibility) than Memphis Medium and Memphis Bold (Luckiesh & Moss, 1940).

4.3.4 SHAW, 1969

Shaw (1969) investigates the legibility of Plantin and Gill Sans in two weights (Figure 29) at two sizes. Participants are 288 partially sighted adults, all over 18 years of age, and 61% over 65 years of age. Using test material printed at sizes near each participant's visual threshold, legibility is measured as reading speed. This study is introduced in Chapter 2 (section 2.4.3.2).

Test material consists of short paragraphs (mixed case) printed in each typeface (Plantin and Gill Sans), at two weights (Roman and Bold), and at six point sizes (12, 14, 16, 18, 20, and 24 point). Instead of paper, a "stiff light board" is used so that the test material remains flat (Shaw, 1969, p.25). Thirty-two unique test passages are created, consisting of "semantically anomalous random sentences" to eliminate, as much as possible, the use of contextual clues (e.g. "Hungry bridges describe expensive farmers") (Shaw, 1969, p.32). Consistency across passages is considered, with each made up of six sentences, across five lines, with thirty-eight to forty-one characters and spaces per line.

Participants' visual acuity is measured and used as the basis for choosing two point sizes nearest to their visual threshold. Shaw (1969, p.24) explains: "If he could only just see 14 point on the ophthalmic test chart, for example, he would be given typographic tests in 12 and 14 point; if he could just see 20 point, the tests would be in 18 and 20 point." Subjects are not tested with each typographic variation, and read a total of four passages each. Experimental design ensures that each subject is tested with two passages in each typeface, weight, and size. For each test reading, the total reading time and the number of words correctly read is recorded. Reading speed is calculated as the average time per correct word.

Shaw (1969) undertakes statistical analysis to assess the influence of the typographic factors on reading speed. The effect of weight and size are statistically significant, and Shaw concludes that "Increased weight or boldness of type, although of secondary importance compared with size, also improves legibility for most partially-sighted [sic] readers.” (Shaw, 1969, p.65). The study also finds that readers with cataracts and glaucoma benefit from increased weight, while those with macular degeneration do not.

Shaw’s (1969) statistical analysis also reveals a significant interaction between size and weight. While larger size and bolder weight is more legible, when the two changes are combined, the smaller bold and larger roman weight perform better than expected, and the larger bold and smaller roman perform worse. Shaw’s (1969) analysis indicates that an increase in the size of a light typeface improves reading speed more than the same size increase of a bold typeface. This suggests that increased stroke width is more important at smaller sizes.

My analysis (Table 4; Figure 30) indicates that the Plantin and Gill Sans typeface families illustrate the typical relationship between increased stroke width and increased letter width (section 2.3.1). While Plantin Bold has a heavier stroke width (13% higher), it also has an increased letter width (14% higher) compared to Plantin Roman. Similarly, Gill Sans Bold has a heavier stroke width (11% higher), as well as an increased letter width (16% higher) compared to Gill Sans Roman. Regarding letter spacing, both typeface families show the typical relationship between increased stroke width and decreased letter spacing (section 2.3.1). The differences in letter spacing between the Bold and Roman typefaces are relatively minor, suggesting little impact on legibility across typeface weights. Plantin Bold has 1% lower letter spacing than Plantin Roman, and Gill Sans Bold has 2% lower letter spacing than Gill Sans Roman (Table 4).

This scientific review makes an assumption that individual bold typefaces are more legible than their regular weight counterparts (Figure 30), although this is not tested directly. While the analysis reveals a significant interaction between typeface and weight, the study notes that “the interactions were of little statistical significance compared with the main effects” (Shaw, 1969, pp.47-48). As these typefaces are all included in Shaw’s (1969) statistical analysis, they are visualised together (Figure 30).

Lastly, it is appropriate to note that Shaw (1969) refers to the lighter weight of Plantin and Gill Sans as Medium weight. However, the test material printed in the study's appendix labels the lighter weight as Roman. The typefaces I measure and visualise (Gill Sans Regular and Plantin Regular) are a match for Shaw's (1969) test material, therefore I refer to these typefaces as Roman.

Table 4. Scientific review parameter data: Shaw, 1969. The table presents typefaces tested experimentally within the study, and their parameter values based on my measurements.

Typeface Tested	**Parameter Values (% x-height)**		
	Stroke Width	**Letter Width**	**Letter Spacing**
Plantin Bold	32	101	38
Plantin Roman	19	87	40
Gill Sans Bold	31	100	26
Gill Sans Roman	20	84	27

Plantin Bold abcdefghijklmnopqrstuvwxyz
Plantin abcdefghijklmnopqrstuvwxyz
Gill Sans Bold abcdefghijklmnopqrstuvwxyz
Gill Sans abcdefghijklmnopqrstuvwxyz

Figure 29: Plantin and Gill Sans, Bold and Roman, tested by Shaw (1969).

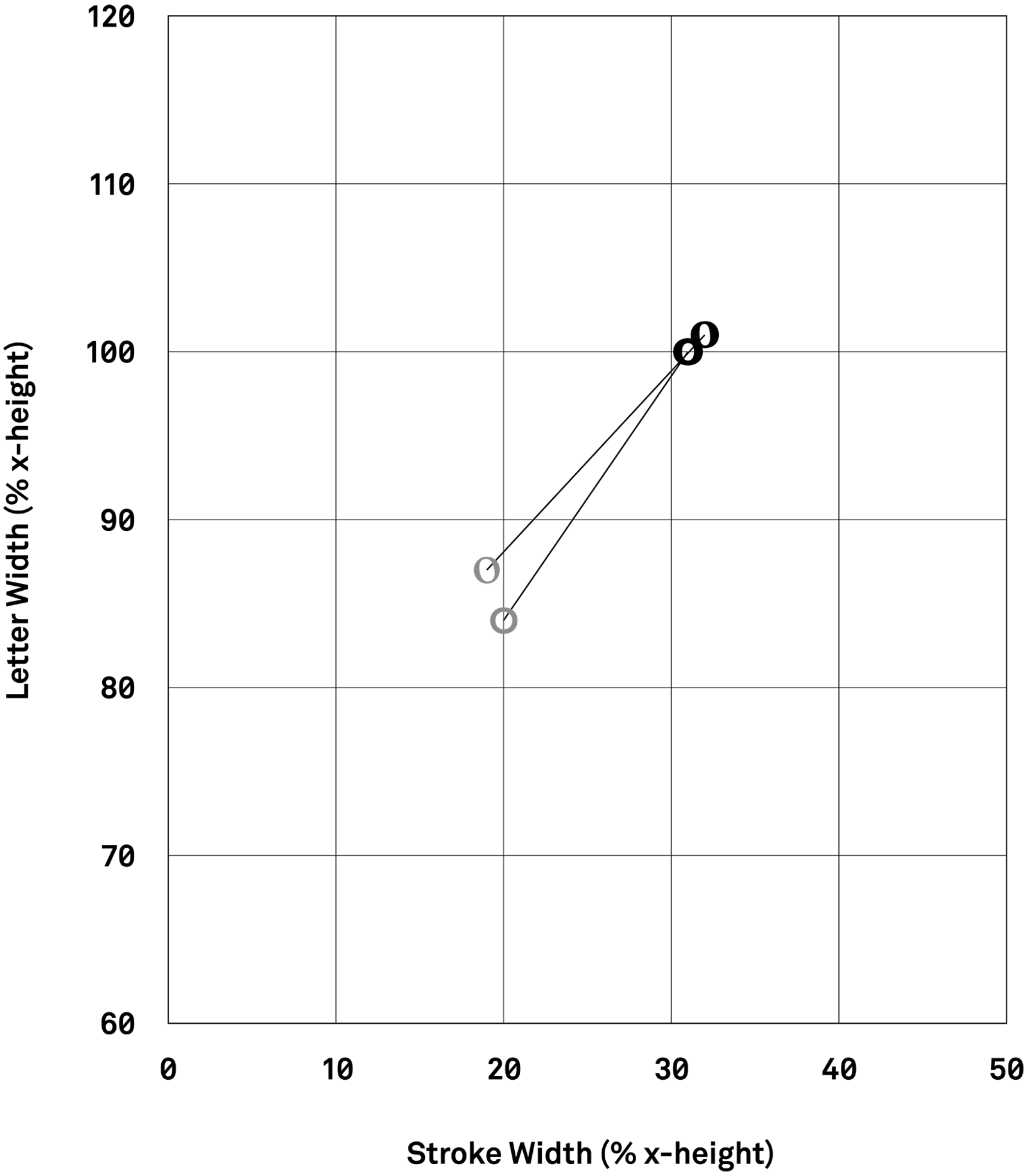


Figure 30: Plantin Roman and Gill Sans Roman are found to be less legible (reading speed) than Plantin Bold and Gill Sans Bold, for partially sighted participants (Shaw, 1969). Typefaces can be identified on the graph because Plantin Bold has a larger stroke width than Gill Sans Bold, and each typeface family is joined by a line.

4.3.5 SMITHER AND BRAUN, 1994

Smither and Braun (1994) investigate the legibility of regular and bold weights of Courier, Century Schoolbook, and Helvetica with older adult participants. The sixteen participants are aged over 65, and include eight males (mean age = 71.25 years, SD = 5.80 years) and eight females (mean age = 70.6 years, SD = 9.56 years). The test material is designed to represent prescription label information, and legibility is measured as reading speed and error rate. The following reports the results of Experiment 2, which uses test material presented on a flat surface (versus on medication bottles), and is therefore more comparable to the other studies in this review. As stated in section 4.2.1, I am only able to find a match for Smither and Braun’s (1994) Helvetica and Helvetica Bold (Figure 31).

The test material consists of prescription labels (mixed case) printed at 9, 12, and 14 point, and placed on a flat cardboard surface. Four different sets of label information are created, with each including: “(1) name of a physician, (2) name of a medication, (3) instructions, and (4) number of refills” (Smither & Braun, 1994, p.152). The labels are “equated for content of information” (Smither & Braun, 1994, p.152). The time taken to read each label is measured using a stopwatch, with accuracy also recorded. Legibility is assessed based on reading speed as well as error rate.

Smither and Braun (1994) find a statistically significant effect of typeface weight on reading speed and performance errors. The authors state that “Roman weight was slower to read than Bold” and that subjects committed “more errors with Roman than with Bold print” (Smither & Braun, 1994, p.156) (Figures 31 and 32). Regarding the statistical analyses, no significant interactions are reported, and therefore this scientific review makes an assumption that individual bold typefaces are more legible than their regular weight counterparts.

According to my analysis (Table 5; Figure 32), Helvetica Bold and Regular illustrate the typical relationship between stroke width and letter width (section 2.3.1). Helvetica Bold has a heavier stroke width than Helvetica Regular (10% higher), and also has an increased letter width (8% higher). While it is common for bolder typefaces to have decreased letter spacing (section 2.3.1), Helvetica Bold and Regular are an example of equivalent letter spacing values across different weights within the same typeface family (Table 5).

Table 5. Scientific review parameter data: Smither & Braun, 1994. The table presents typefaces tested experimentally within the study, and their parameter values based on my measurements.

Typeface Tested	**Parameter Values (% x-height)**		
	Stroke Width	**Letter Width**	**Letter Spacing**
Helvetica Bold	27	90	25
Helvetica	17	82	25

Helvetica Bold abcdefghijklmnopqrstuvwxyz

Helvetica abcdefghijklmnopqrstuvwxyz

Figure 31: Helvetica Bold and Regular, tested by Smither and Braun (1994).

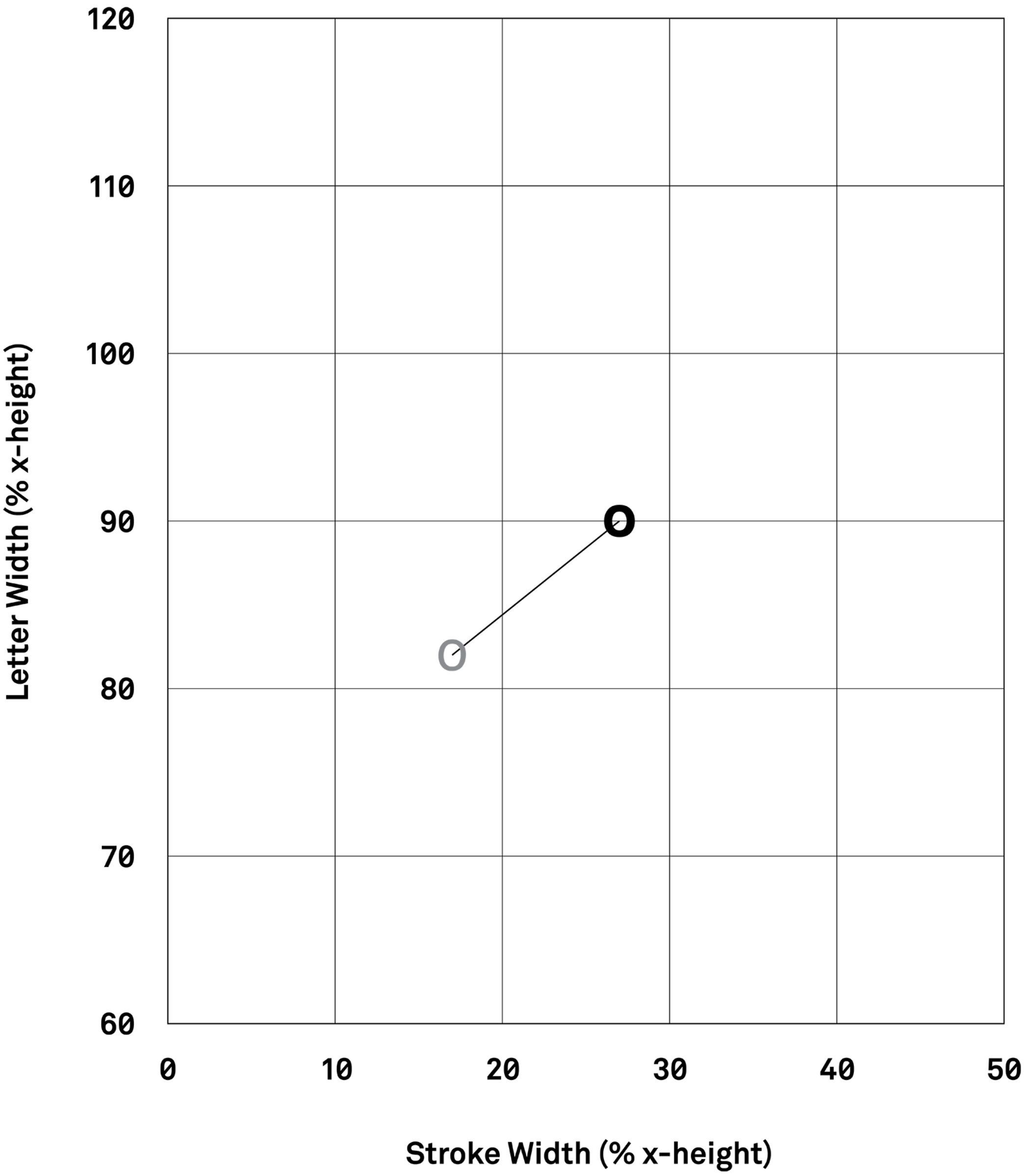


Figure 32: Helvetica is found to be less legible (reading speed, performance errors) than Helvetica Bold for older participants (Smither & Braun, 1994).

4.3.6 MANSFIELD, LEGGE, AND BANE, 1996

Mansfield et al. (1996) investigate the legibility of Times Roman and Courier Bold (Figure 14). The forty-two low vision participants (twenty with AMD) are aged 23-83, and include subjects with intact central vision (mean age = 41 years, SD = 11.2 years) and central vision loss (mean age = 68 years, SD = 15.4 years). Reading performance is assessed using two versions of the MNREAD Acuity Chart (Mansfield et al., 1993) (section 2.2.2), printed with Times Roman and Courier Bold. The study by Mansfield et al. (1996) is introduced in Chapter 2 (section 2.4.3.5).

The MNREAD Acuity Chart presents a series of nineteen sentences (mixed case) printed at progressively smaller sizes. The sentences are comprised of the most common words in printed English and are matched for reading difficulty. The Times Roman sentences have sixty characters (including spaces) and are printed onto three lines of justified text. The Courier Bold sentences have fifty-six characters and are printed onto four lines of text. The two typefaces are scaled to achieve equivalent x-height. The charts are placed on a reading stand in front of the subject who reads each sentence starting with the largest print size. Subjects continue reading until they cannot read any words in a sentence. The time taken to read each sentence and any reading errors are recorded. The MNREAD Acuity Chart is used to determine reading acuity and CPS for each participant. Reading acuity is the smallest print that can just be read, and CPS is the smallest print that can be read at the maximum reading speed. Reading speed is measured in words per minute (wpm) and determined for each sentence as the number of standard-length words read correctly, divided by the time taken to read the sentence.

Mansfield et al. (1996) find that Times Roman is less legible than Courier Bold (Figures 14 and 33). Reading acuity scores obtained with Times Roman are significantly poorer than with Courier Bold. CPS is significantly larger with Times Roman than with Courier Bold. As quoted in Chapter 2, the authors acknowledge that “Any of the differences between the fonts might be expected to influence reading performance.” (Mansfield et al.,1996, p.1493). While the bold weight of Courier has a slightly heavier stroke width (2% higher) than Times Roman, it is also a monospaced typeface and has notably increased letter width (16% higher) and letter spacing (11% higher) (Table 6; Figures 14 and 33). Based on these magnitude differences, my calculations suggest that the improved legibility of Courier Bold may be attributed to these parameters associated with monospaced typefaces (i.e. letter width, letter spacing).

Table 6. Scientific review parameter data: Mansfield et al., 1996. The table presents typefaces tested experimentally within the study, and their parameter values based on my measurements.

Typeface Tested	**Parameter Values (% x-height)**		
	Stroke Width	**Letter Width**	**Letter Spacing**
Courier Bold	21	92	46
Times Roman	19	76	35

Courier Bold abcdefghijklmnopqrstuvwxyz
Times Roman abcdefghijklmnopqrstuvwxyz

Figure 14: Courier Bold and Times Roman, tested by Mansfield et al. (1996). Figure reproduced from Chapter 2.

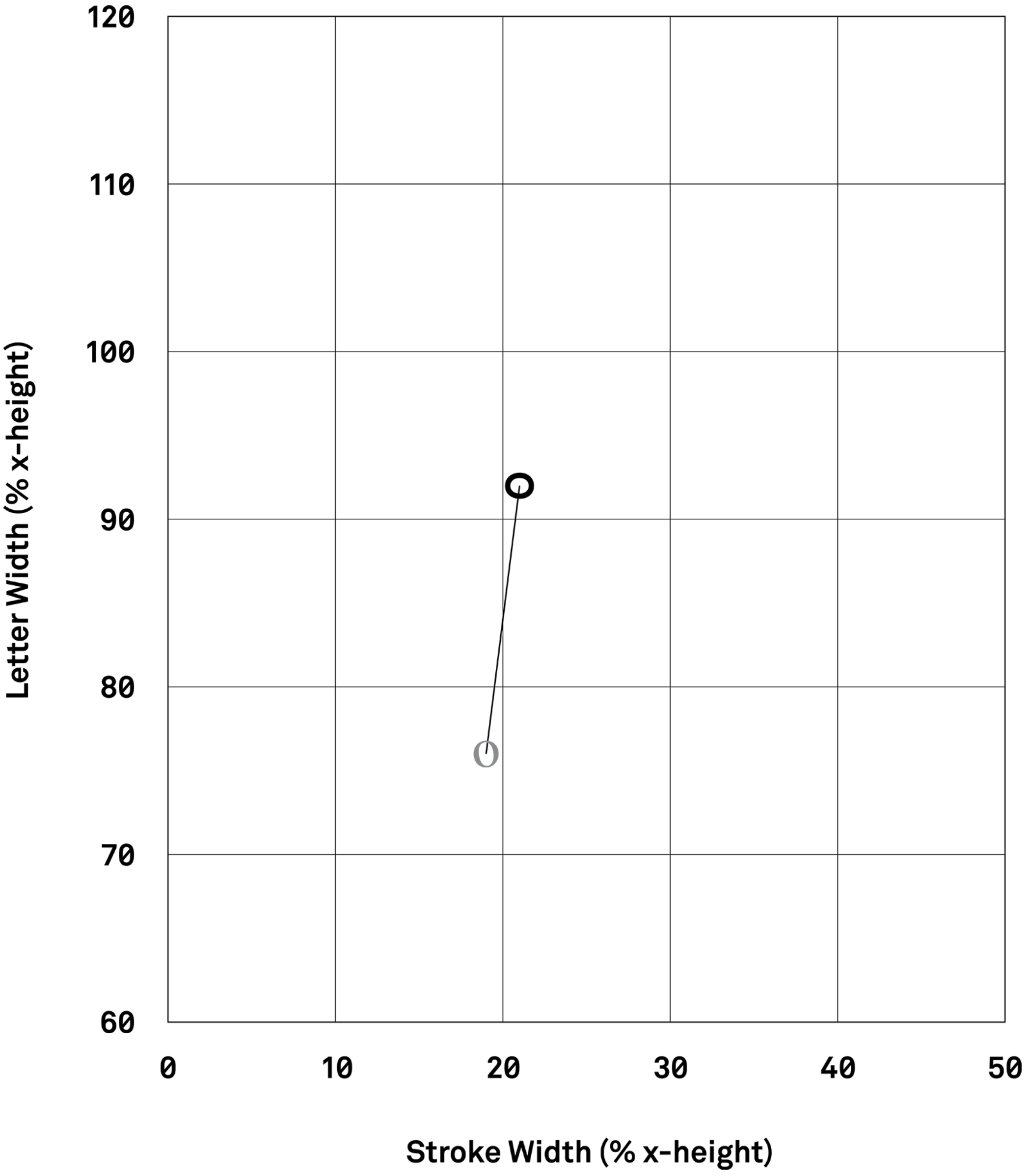


Figure 33: Times Roman is found to be less legible (reading acuity, CPS) than Courier Bold for low vision participants (Mansfield et al., 1996).

4.3.7 SHEEDY, SUBBARAM, ZIMMERMAN, AND HAYES, 2005

Sheedy et al. (2005) conduct two experiments which investigate the influence of stroke width on legibility, measured as visual acuity. Thirty normally sighted participants are aged 18-35. The results from Experiment 1 and Experiment 4 are presented here. Experiment 1 tests the regular and bold weights of Georgia, Verdana, Times New Roman, and Arial (Figure 34), and Experiment 4 tests four weights of Franklin Gothic (Figure 4). Sheedy et al. (2005) is introduced in Chapter 2 (section 2.3.2).

In Experiment 1, the test material is presented as uppercase letters and lowercase words on screen (LCD, CRT) and in print. The test material consists of five letters or five words (five to six letters per word), set in each typeface (regular and bold weights of Georgia, Verdana, Times New Roman, and Arial), at four point sizes (8, 10, 12, and 14 point). All combinations of variables are not tested with each participant.

Instead of presenting test material at progressively smaller sizes to measure visual acuity, letter size is held constant and viewing distance is altered. This is done to avoid the resolution issues of small type sizes on screen. Sheedy et al.'s (2005) procedure entails testing sequentially greater distances until the participant cannot correctly identify any of the letters or words. Visual acuity is calculated based on the number of correct letters or words identified. Sheedy et al. (2005) also calibrate the data to address the discrepancy between point size and vertical letter size across the typefaces tested.

Statistical analysis indicates a significant effect of bold weight on the legibility of both uppercase letters and lowercase words. Regarding the legibility of bold weights, Sheedy et al. (2005, p.803) state: “Bold letters have wider stroke widths, but the entire character is also wider; increased legibility could be attributable to one or both factors.” My analysis (Table 7; Figure 35) demonstrates this typical relationship between stroke width and letter width (section 2.3.1). While Georgia, Verdana, Times New Roman, and Arial Bold have heavier stroke widths than their Regular counterparts (14%, 15%, 12%, and 9% higher, respectively), they also have increased letter widths (22%, 16%, 20%, and 10% higher, respectively). The analysis also illustrates that when studies test regular and bold typefaces, there is notable diversity in proportions within both groups.

My analysis (Table 7) also demonstrates the typical relationship between stroke width and letter spacing (section 2.3.1), in three of the four typeface families investigated. Georgia, Verdana, and Times New Roman Bold have decreased letter spacing compared to their Regular counterparts (2%, 3%, and 20% lower, respectively). Arial Bold is an exception, with increased letter spacing (1% higher) compared to Regular. While letter spacing would not affect the legibility of uppercase letters, it should influence the legibility of lowercase words. These calculations suggest that legibility is improved by increased stroke width and/or letter width, despite decreases in letter spacing in three of the four bold typefaces.

This review makes an assumption that individual bold typefaces are more legible than their regular weight counterparts (Figure 35), as no interactions are reported between weight and typeface. The four typeface families are all included in Sheedy et al.’s (2005) analysis, therefore they are visualised together (Figure 35). Note that the authors refer to “Bold” as being “On/off” (Sheedy et al., 2005, p.801), and no test material is presented within the publication, therefore I measure and visualise system fonts.

Table 7. Scientific review parameter data: Sheedy et al., 2005 (Experiment #1). The table presents typefaces tested experimentally within the study, and their parameter values based on my measurements.

Typeface Tested	Parameter Values (% x-height)		
	Stroke Width	Letter Width	Letter Spacing
Georgia Bold	33	103	39
Georgia	19	81	41
Verdana Bold	32	100	29
Verdana	17	84	32
Times New Roman Bold	30	97	14
Times New Roman	18	77	34
Arial Bold	26	91	27
Arial	17	81	26

Georgia Bold **abcdefghijklmnopqrstuvwxyz**
Georgia abcdefghijklmnopqrstuvwxyz
Verdana Bold **abcdefghijklmnopqrstuvwxyz**
Verdana abcdefghijklmnopqrstuvwxyz
Times New Roman Bold **abcdefghijklmnopqrstuvwxyz**
Times New Roman abcdefghijklmnopqrstuvwxyz
Arial Bold **abcdefghijklmnopqrstuvwxyz**
Arial abcdefghijklmnopqrstuvwxyz

Figure 34: Georgia Bold, Verdana Bold, Times New Roman Bold, and Arial Bold, and their Regular counterparts, tested by Sheedy et al. (2005).

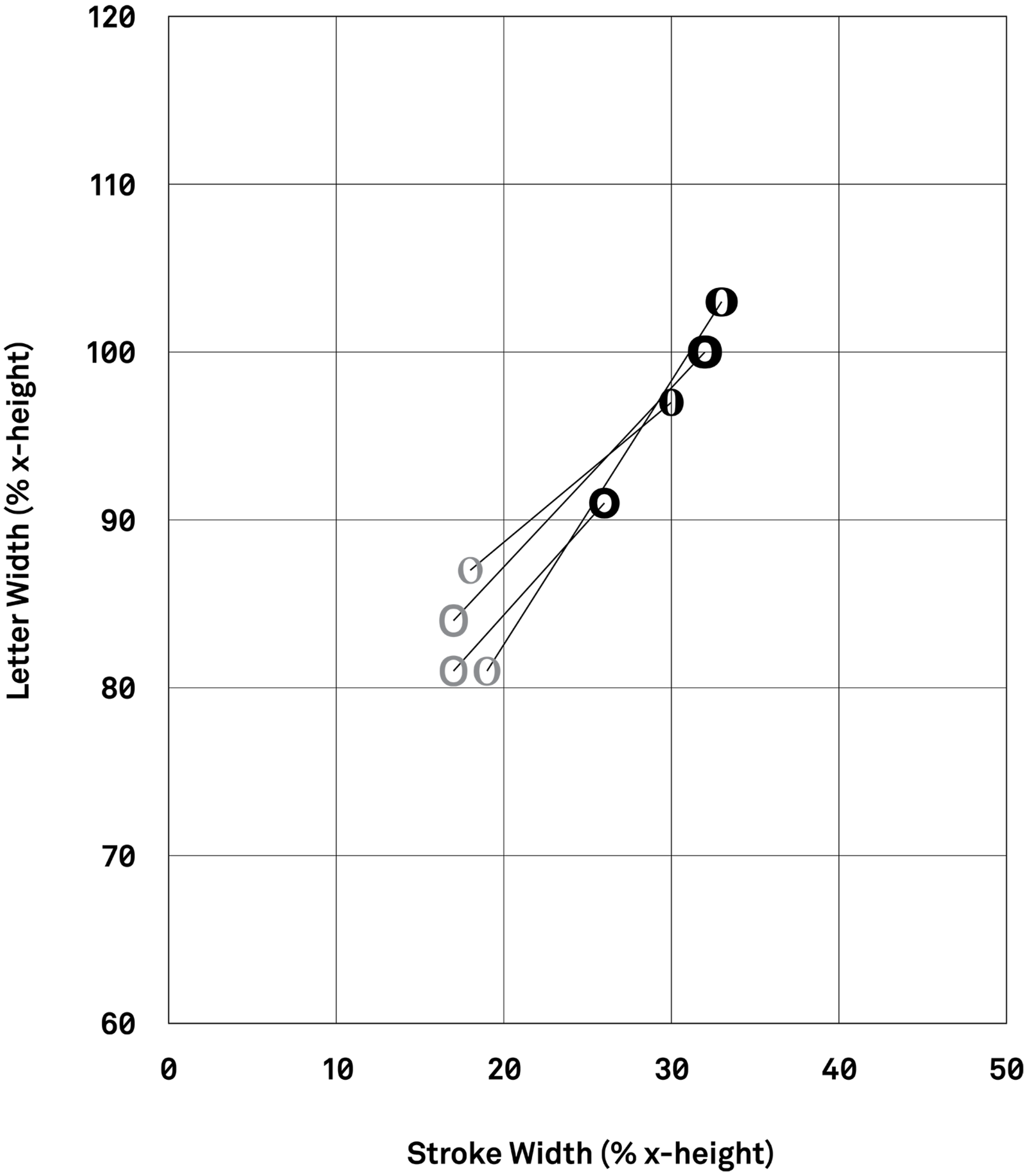


Figure 35: Arial, Verdana, Times New Roman, and Georgia Regular are found to be less legible (visual acuity) than Arial, Verdana, Times New Roman, and Georgia Bold (Sheedy et al., 2005). Typefaces can be identified on the graph because Georgia Bold has the largest stroke width value, followed by Verdana Bold, Times New Roman Bold, and Arial Bold. Regular typefaces can be identified by the line joining typeface families.

Experiment 4 investigates four weights of Franklin Gothic: Book, Medium, Demi, and Heavy (Figure 4). Test material is presented as uppercase letters, lowercase letters, and lowercase words at 12 point on screen (LCD). As in Experiment 1, the test material consists of five letters or five words (five to six letters per word), and all combinations of variables are not tested with each participant. Experiment 4 also employs the same methods to measure visual acuity, holding letter size constant and altering viewing distance.

Sheedy et al. (2005) find that Franklin Gothic Book is significantly less legible than Franklin Gothic Medium, Demi, and Heavy (Figures 4 and 36). However, there is no significant difference among the three heavier weights (Sheedy et al., 2005). More detailed results are reported in the PhD thesis of the second author of this study (Subbaram, 2004). Subbaram (2004) reports significant interaction effects. Explaining these results he states: “The relative legibility of the upper case [sic] letters increases initially with increased stroke width but decreases for the heaviest stroke width.” and “relative legibility of the lower case [sic] letters and words increase with increased stroke width and attain greater legibility at the heaviest stroke width” (Subbaram, 2004, p.143). Summarising these findings, Subbaram states that “for the upper case [sic] letters, there is a decrease in relative legibility at the heaviest stroke width” (Subbaram, 2004, p.144) and “lower case [sic] letters and words were more legible at the heaviest stroke widths” (Subbaram, 2004, p.145).

Similar to the analysis of typefaces tested in Experiment 1, my analysis of Franklin Gothic (Table 8; Figure 36) indicates that increased stroke width is accompanied by increased letter width. While Franklin Gothic Book, Medium, Demi, and Heavy have progressively larger stroke width values (15%, 22%, 29%, and 39%, respectively), they also have larger letter width values (79%, 83%, 87%, and 98%, respectively). My analysis of Franklin Gothic also reveals decreased letter spacing in the heavier weights. While Franklin Gothic Book and Medium have equivalent letter spacing (27%), letter spacing is decreased in the Demi (20%) and Heavy (16%) weights. This letter spacing data suggests that lowercase word legibility is improved by increased stroke width and/or letter width, despite decreases in letter spacing at the Demi and Heavy weights.

Table 8. Scientific review parameter data: Sheedy et al., 2005 (Experiment #4). The table presents typefaces tested experimentally within the study, and their parameter values based on my measurements.

Typeface Tested	**Parameter Values (% x-height)**		
	Stroke Width	**Letter Width**	**Letter Spacing**
Franklin Gothic Heavy	39	98	16
Franklin Gothic Demi	29	87	20
Franklin Gothic Medium	22	83	27
Franklin Gothic Book	15	79	27

Franklin Gothic Heavy abcdefghijklmnopqrstuvwxyz
Franklin Gothic Demi abcdefghijklmnopqrstuvwxyz
Franklin Gothic Medium abcdefghijklmnopqrstuvwxyz
Franklin Gothic Book abcdefghijklmnopqrstuvwxyz

Figure 4: Franklin Gothic Heavy, Demi, Medium, and Book, tested by Sheedy et al. (2005). Figure reproduced from Chapter 2.

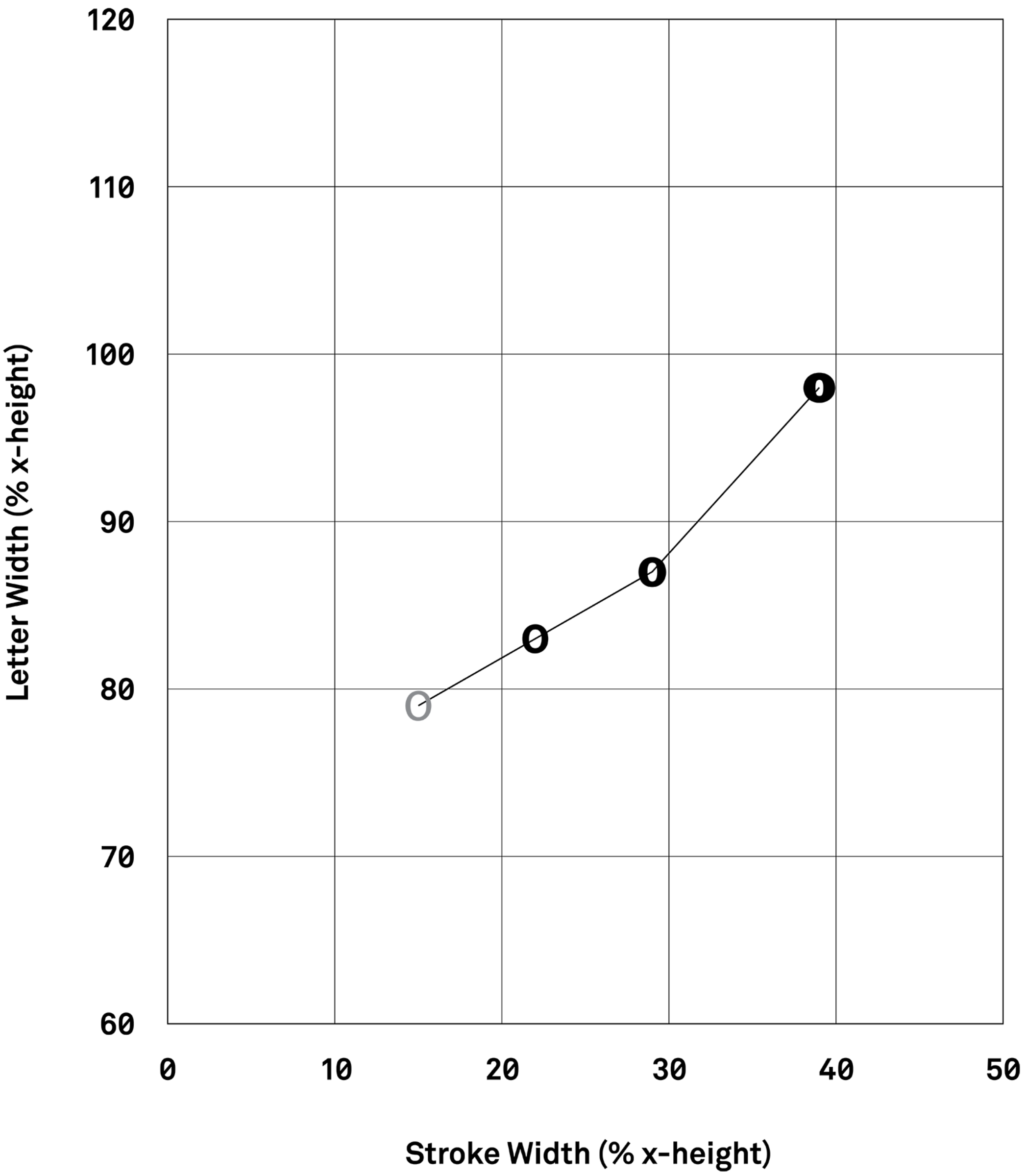


Figure 36: Franklin Gothic Book is found to be less legible (visual acuity) than Franklin Gothic Medium, Demi, and Heavy (Sheedy et al., 2005).

4.3.8 BEVERATOU, 2016

Beveratou (2016) investigates the legibility of ten typefaces (Appendix 3c). Participants are twenty-one partially sighted adults (ten with macular degeneration) aged 70-94. Legibility is measured as reading speed. This study is introduced in Chapter 2 (section 2.5.2).

The test material consists of paragraphs (lowercase) set in each of the ten typefaces. Each typeface is scaled to have an x-height equivalent to Arial 16 point. The paragraphs consist of eighty words (varying from two to eight letters), placed in random order. The participants read each paragraph aloud for twenty seconds, and the last word read is recorded. Legibility is measured as reading speed; the number of words read in twenty seconds. The total number of words read by the twenty-one participants for each of the typefaces is reproduced from Beveratou's (2016) study in Appendix 3c. No statistical analyses are reported.

Based on the total number of words read for each typeface (Appendix 3c), I analyse Tiresias LPfont (most legible), Freight Sans Book (least legible), and Arial Regular (typeface from the mid-range) (Table 9; Figures 37 and 38). Arial Regular is chosen because it is the only typeface from the mid-range with low stroke contrast, which is the most appropriate choice for a review focused on stroke width. Regarding Tiresias LPfont, Beveratou states that its "success may be due to the thickness of the font" (Beveratou, 2016, p.12). My analysis (Table 9; Figure 38) illustrates that Tiresias LPfont has a larger stroke width value (26%), compared to Arial Regular (17%), and Freight Sans Book (13%). My analysis also illustrates increased letter width with increased stroke width, with the letter width values of the three typefaces within a relatively small range: Tiresias LPfont (83%), Arial Regular (81%), and Freight Sans Book (79%). Letter spacing values are also within a relatively small range: Tiresias LPfont (29%), Arial Regular (26%), and Freight Sans Book (31%). Based on the magnitude differences in parameter values across the typefaces, my data—in agreement with Beveratou's statement—suggests that the performance benefit of Tiresias LPfont may be due to its increased stroke width.

Table 9. Scientific review parameter data: Beveratou, 2016. The table presents typefaces tested experimentally within the study, and their parameter values based on my measurements.

Typeface Tested	**Parameter Values (% x-height)**		
	Stroke Width	**Letter Width**	**Letter Spacing**
Tiresias LPfont	26	83	29
Arial Regular	17	81	26
Freight Sans Book	13	79	31

Tiresias LPfont	**abcdefghijklmnopqrstuvwxyz**
Arial	abcdefghijklmnopqrstuvwxyz
Freight Sans Book	abcdefghijklmnopqrstuvwxyz

Figure 37: Tiresias LPfont, Arial Regular, and Freight Sans Book, tested by Beveratou (2016).

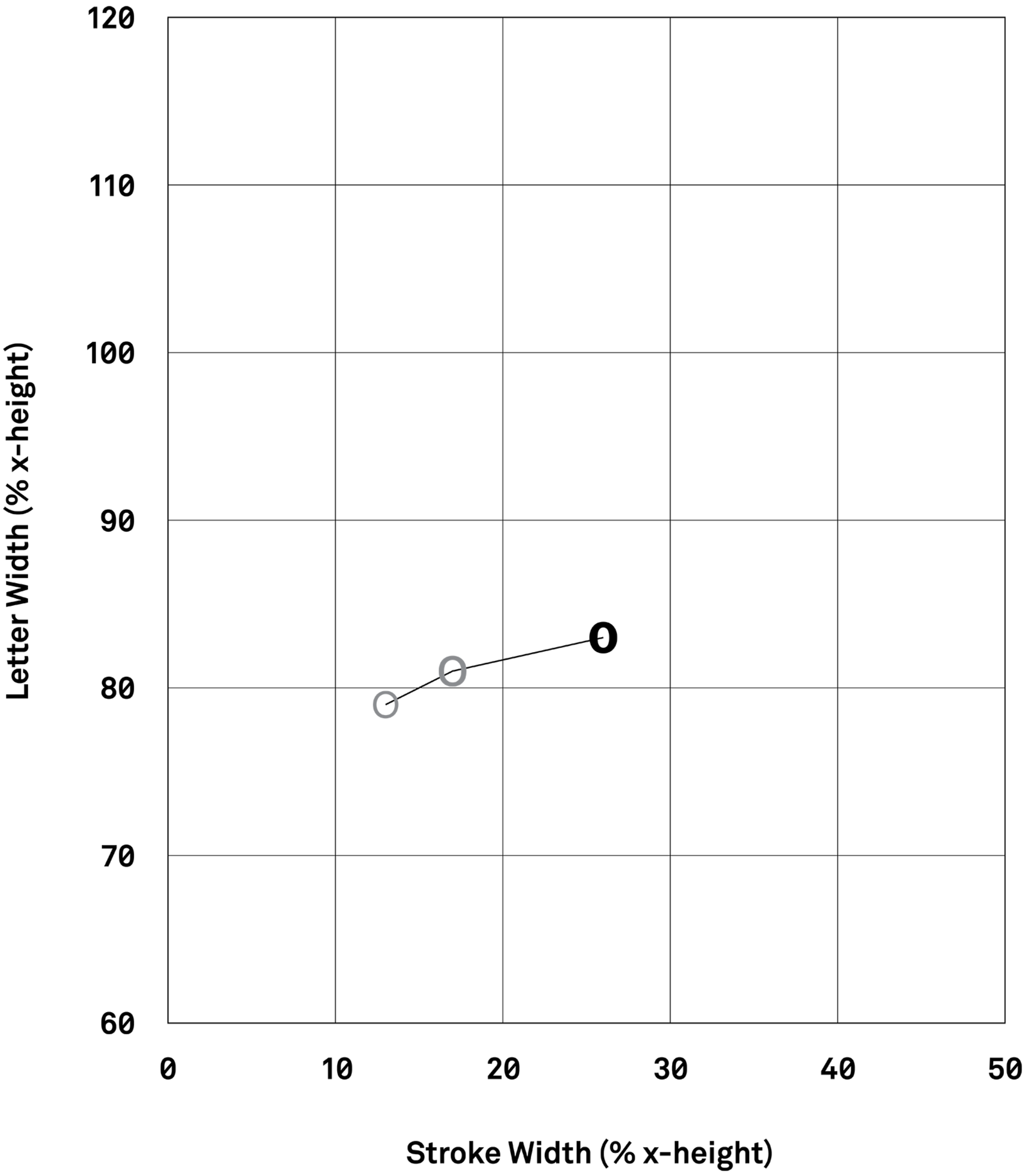


Figure 38: Freight Sans Book and Arial Regular are found to be less legible (reading speed) than Tiresias LPfont for partially sighted participants (Beveratou, 2016).

4.3.9 XIONG, LORSUNG, MANSFIELD, BIGELOW, AND LEGGE, 2018

Xiong et al. (2018) investigate the legibility of Maxular Rx Bold, Courier, Helvetica, and Times (Figure 39) with low vision participants. Reading performance is tested with nineteen subjects diagnosed with macular degeneration (mean age = 65.0 years, SD = 10.03 years). Reading performance is evaluated with digital versions of the MNREAD test rendered with each of the typefaces. Xiong et al. (2018) also test the typeface Eido (section 2.5.3), however because it has unconventional letterforms it is not included in this analysis. The study by Xiong et al. (2018) and Eido are introduced in Chapter 2 (section 2.5.3).

Different from the chart-based test used by Mansfield et al. (1996) (section 4.3.6), Xiong et al. (2018) display MNREAD sentences on screen (17-inch MacBook Pro laptop computer). Each typeface is tested with fourteen different sentences (mixed case), which are identical across the typefaces. The typefaces are scaled to achieve equivalent x-height. As in Mansfield et al. (1996), the MNREAD test is used to determine reading acuity and CPS for each participant. Reading acuity is the smallest print size that can just be read, and CPS is the smallest print size yielding the best reading speed.

Xiong et al. (2018) find that Maxular Rx Bold has a significantly better reading acuity than Helvetica and Times, but not Courier (Figures 39 and 40). Maxular Rx Bold also permits significantly smaller CPS than Helvetica and Times, but again does not show an advantage over Courier. Courier has the smallest mean CPS, which is significantly smaller than Helvetica.

Maxular Rx is designed specifically for people with macular degeneration (section 2.5.3). This typeface is “designed to be very bold” (Xiong et al., 2018, p.4182) with increased letter and word spacing (Delve Fonts, 2021). While Maxular Rx Bold has a larger stroke width value (11% and 9% higher compared to Helvetica and Times, respectively), it also has a much larger letter width (25% and 30% higher than Helvetica and Times, respectively)

(Table 10; Figures 39 and 40). Further calculations show that Maxular Rx Bold also has notably increased letter spacing (53% and 43% higher compared to Helvetica and Times, respectively) (Table 10). The differences in stroke width between Maxular Rx Bold versus Helvetica and Times are much less dramatic than the differences in letter width and letter spacing. As introduced in Chapter 3 (section 3.4.5), Xiong et al. (2018) investigate the differences in spacing between the typefaces they test. Their analysis reveals that typefaces with larger spacing permit smaller reading acuity and CPS (Xiong et al., 2018). In contrast to the methods employed within this PhD, Xiong et al. (2018) calculate typeface spacing as the average centre-to-centre separation between adjacent letters (as a proportion of x-height). As such, their measurement of spacing includes both letter width and letter spacing. Similar to the discussion of Courier Bold (section 4.3.6), this suggests that the improved performance of Maxular Rx Bold is more influenced by increased letter width and spacing versus stroke width.

As stated, Courier is found to have a significantly smaller CPS than Helvetica. Comparing the stroke width and letter width of Courier and Helvetica, they are very similar (differ by 1% in both parameters) (Table 10; Figures 39 and 40). In contrast, Courier has notably increased letter spacing (25% higher) as compared to Helvetica (Table 10). While letter spacing may account for the legibility difference, letter width could also play a role because Courier is a monospaced typeface and therefore generally has larger letter widths across its characters compared to a proportionally spaced typeface. Courier is not included in subsequent analyses because it illustrates a significant difference in legibility between a monospaced versus proportional typeface, rather than typefaces that vary in stroke width which is the focus of the review.

Table 10. Scientific review parameter data: Xiong et al., 2018. The table presents typefaces tested experimentally within the study, and their parameter values based on my measurements.

Typeface Tested	**Parameter Values (% x-height)**		
	Stroke Width	**Letter Width**	**Letter Spacing**
Maxular Rx Bold	28	107	78
Courier	16	83	50
Helvetica	17	82	25
Times Roman	19	77	35

Maxular Rx Bold	**abcdefghijklmnopqrstuvwxyz**
Courier	abcdefghijklmnopqrstuvwxyz
Helvetica	abcdefghijklmnopqrstuvwxyz
Times	abcdefghijklmnopqrstuvwxyz

Figure 39: Maxular Rx Bold, Courier, Helvetica, and Times, tested by Xiong et al. (2018).

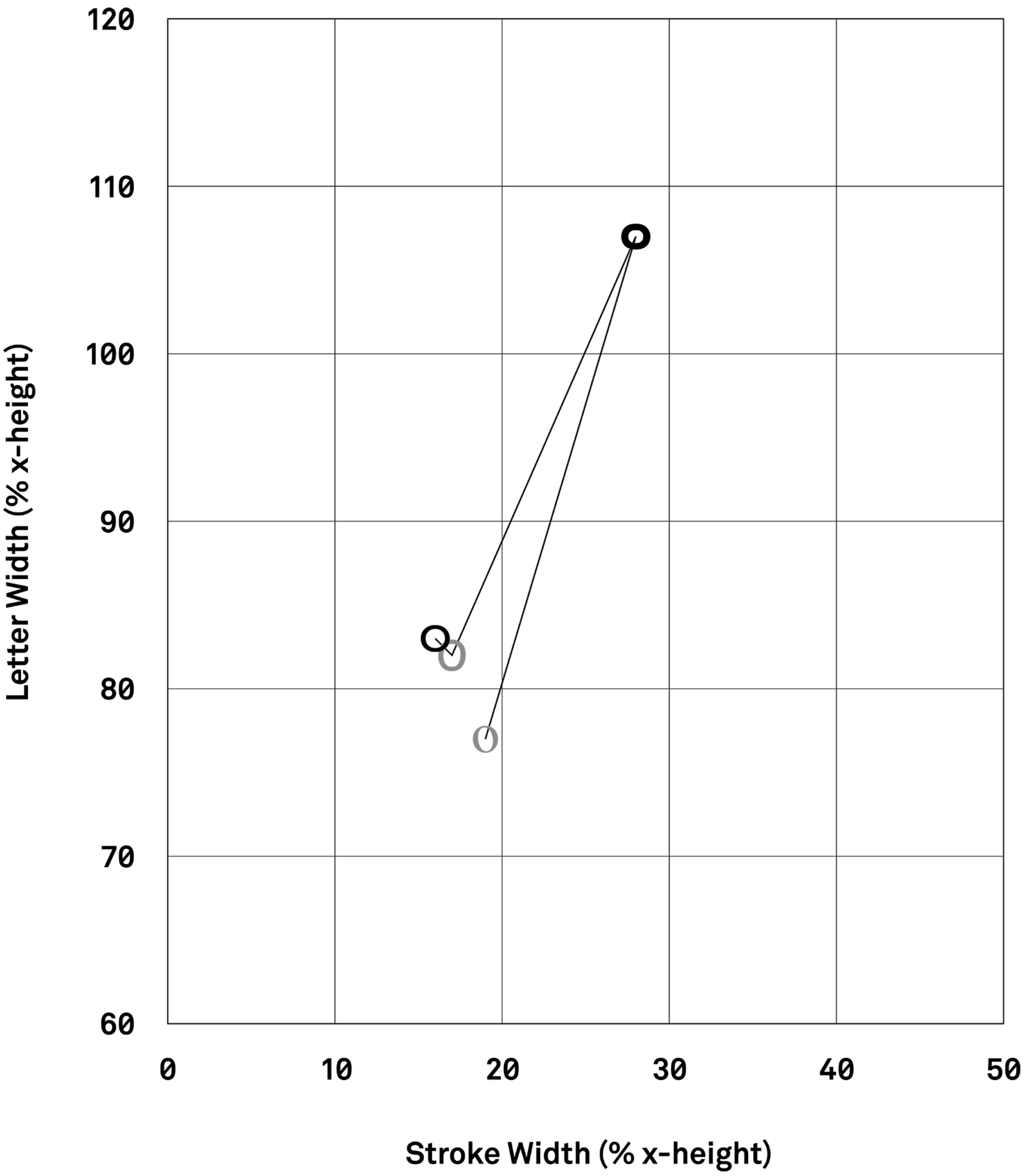


Figure 40: Helvetica and Times are found to be less legible (reading acuity, CPS) (represented by two long lines) than Maxular Rx Bold, and Helvetica is found to be less legible (CPS) (represented by one short line) than Courier, for participants with macular degeneration (Xiong et al., 2018).

4.3.10 BEIER AND ODERKERK, 2019A

Beier and Oderkerk (2019a) investigate the influence of stroke width on legibility with a laboratory typeface family called Ovink (Figure 41). Results of Experiment 2 are presented, which demonstrate a legibility benefit of stroke width beyond regular weight, at a small type size. Experiment 2 tests fifteen normally sighted participants, which include four males and eleven females (mean age = 27.33 years, SD = 5.09 years). Legibility is measured as letter recognition.

The legibility of Ovink Regular, Semi Bold, and Ultra Black are investigated (Figure 41). The laboratory typefaces are designed for the experiment, allowing for the testing of stroke widths that meet the goals of the research. Stroke widths are calculated as a ratio to ascender height (e.g. ‘h’): Ovink Regular (1:10.0), Semi Bold (1:4.7), and Ultra Black (1:2.5) (Figure 41).

The test material is presented as individual lowercase letters at ~6 point. Sixteen different letters are tested (Figure 41). The test material is presented for a short exposure time on screen (CRT), after which participants attempt to identify the letter. Legibility is measured as mean accuracy of the responses for each of the typeface weights.

Statistical analyses show a significantly lower mean accuracy for Ovink Regular than Ovink Semi Bold (Figures 41 and 42). Mean accuracy of Ovink Ultra Black is significantly lower than both Ovink Regular and Ovink Semi Bold (Figures 41 and 42). Only the latter relationship is visualised, illustrating the performance decline of Ovink Ultra Black compared to Semi Bold (the most legible weight). This study demonstrates that letter recognition is enhanced by letter boldness (i.e. Ovink Regular versus Ovink Semi Bold), however performance declines with extreme boldness (i.e. Ovink Ultra Black).

My analysis (Table 11, Figure 42) presents the proportions of Ovink Regular, Semi Bold, and Ultra Black as percentages of x-height. While Beier and

Oderkerk (2019a) present stroke width quantitatively (as a ratio to ascender height), it is useful to present stroke width as a percentage of x-height for two reasons. First, my calculation of stroke width is aligned with design practice (Bigelow, 2019) (section 4.2.2). Second, presenting stroke width as a percentage of x-height allows for comparisons across the ten studies analysed in this scientific review. My analysis also reveals the relationship between typeface stroke width and letter width across the three weights of Ovink. Note that letter spacing is not included in the analysis because Beier and Oderkerk (2019a) present letters individually.

Based on my analysis (Table 11; Figure 42), the stroke width values of the Ovink typefaces are: Regular (14%), Semi Bold (29%), and Ultra Black (53%). Increased stroke width coincides with increased letter width values: Ovink Regular (83%), Semi Bold (97%), and Ultra Black (111%). My calculations indicate that the stroke width of Ovink Ultra Black (53%) is extreme, and when visualised is literally off the chart (Figure 42). Beier and Oderkerk include Ovink Ultra Black in order to investigate the performance of “extreme letter boldness” (2019a, p.4), and describe it as being “as heavy as possible without the letter counters closing up” (2019a, p.5). As the stroke width of Ovink Ultra Black is not representative of typefaces commonly used in design practice, it is removed from subsequent visualisations of scientific data.

Table 11. Scientific review parameter data: Beier & Oderkerk, 2019a. The table presents typefaces tested experimentally within the study, and their parameter values based on my measurements.

Typeface Tested	**Parameter Values (% x-height)**		
	Stroke Width	**Letter Width**	**Letter Spacing**
Ovink Ultra Black	53	111	-
Ovink Semi Bold	29	97	-
Ovink Regular	14	83	-

Redacted

Figure 41: Ovink Regular (2), Semi Bold (4), and Ultra Black (6), tested by Beier and Oderkerk (2019a). Ratios on the right side of the figure represent the stroke width to ascender height (e.g. ‘h’) ratio of each typeface (see Beier & Oderkerk, 2019a). This article was published in *Acta Psychologica*, 199, Beier, S. and Oderkerk, C.A.T., Smaller Visual Angles Show Greater Benefit of Letter Boldness Than Larger Visual Angles, pp.1-8, Copyright Elsevier (2019).

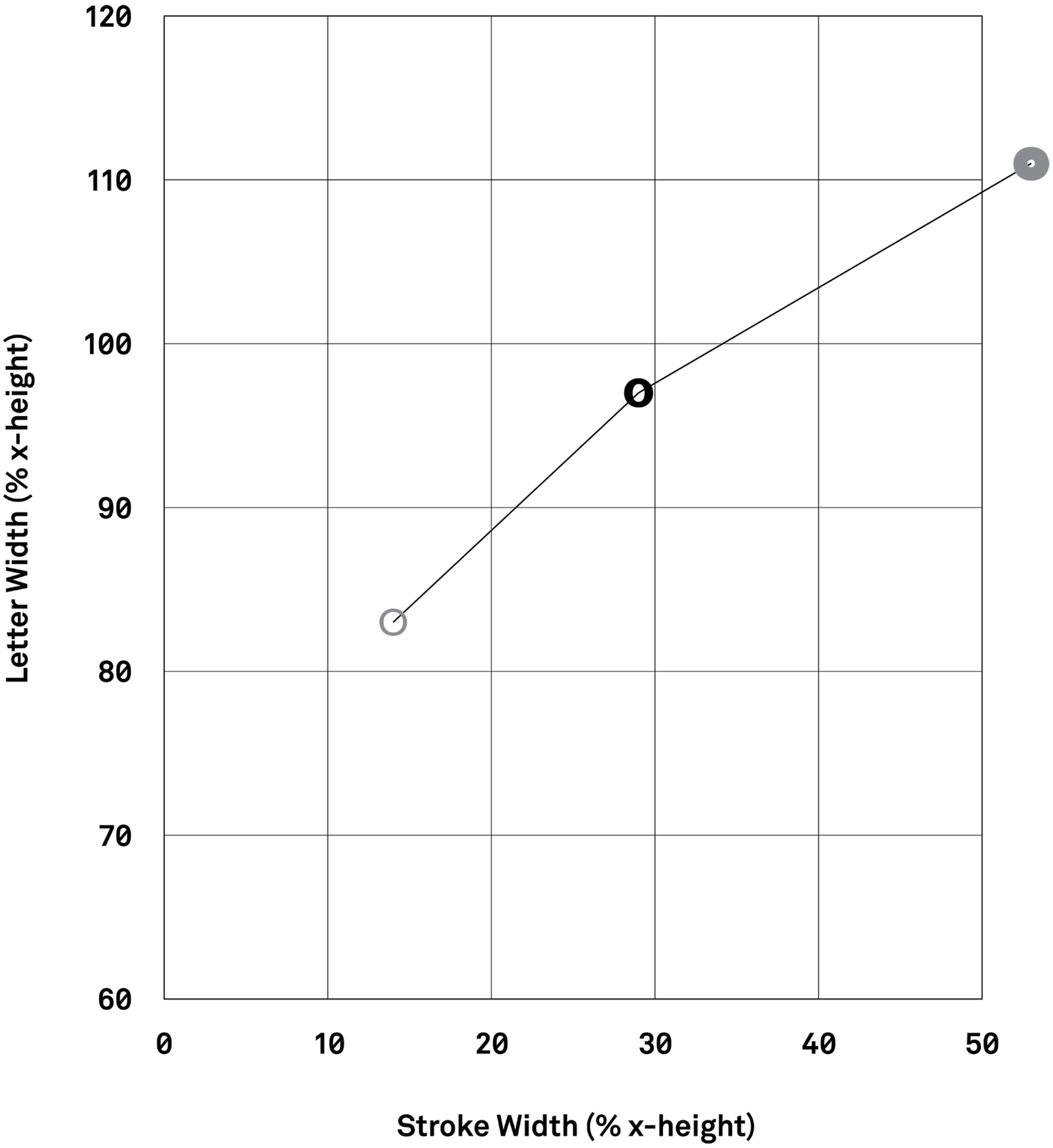


Figure 42: Ovink Regular and Ovink Ultra Black are found to be less legible (letter recognition) than Ovink Semi Bold (Beier & Oderkerk, 2019a). The Ovink Typeface family is represented by ‘o’s with similar proportions (section 4.2.3).

4.3.11 BEIER AND ODERKERK, 2019B

Beier and Oderkerk (2019b) investigate the legibility of Gill Sans Light and KBH Text Regular (Figure 43). The typefaces are tested with twenty participants over the age of 50, consisting of twelve males and eight females (mean age = 67.36 years, SD = 10.3 years). Reading performance is assessed using test material based on the Radner Reading Chart (Radner, 2017), with text printed in both typefaces. Beier and Oderkerk (2019b) also test another typeface—KBH Display Regular—however it is not included in this review as the main difference between KBH Text Regular and KBH Display Regular is in the design of the characters versus proportions (Figure 43).

Beier and Oderkerk's (2019b) test material is based on twenty-eight sentences from the Radner Reading Chart, printed at progressively smaller sizes. The sentences (mixed case) each contain fourteen words, eighty-two to eighty-four characters (including spaces), and are presented over three lines (Radner, 2017). Figure 43 shows sentences from the Danish version of the Radner Reading Chart set in the typefaces tested in the study. The typefaces are scaled to achieve equivalent x-height.

Legibility is measured as reading acuity and CPS. The procedure entails participants reading the sentences consecutively, starting from the largest print size. Reading acuity is based on the smallest sentence of which more than 80% is read correctly. Reading speed for each sentence is measured in words per minute (wpm), determined by the number of correctly read words and the time taken to read the sentence. CPS is based on the smallest sentence read at the maximum reading speed.

Statistical analyses indicate that reading acuity and CPS scores are significantly better for KBH Text Regular than Gill Sans Light. While the authors acknowledge that “when fonts of different families are tested, the fonts vary on several parameters” (Beier & Oderkerk, 2019b, p.59), they suggest that the increased legibility of KBH Text Regular is related to its increased stroke width, letter width, and letter spacing. My analysis (Table 12; Figure 44) indicates that KBH Text Regular has a heavier stroke width (5% higher), an increased letter width (12% higher), and increased letter spacing (13% higher). It is noteworthy that KBH Text Regular has a larger letter width than other typefaces with similar stroke width values (see Figures 44 and 45).

Table 12. Scientific review parameter data: Beier & Oderkerk, 2019b. The table presents typefaces tested experimentally within the study, and their parameter values based on my measurements.

Typeface Tested	**Parameter Values (% x-height)**		
	Stroke Width	**Letter Width**	**Letter Spacing**
KBH Text Regular	16	93	41
Gill Sans Light	11	81	28

KBH Display Regular

Min morfar havde også tømt
sit handskerum, hvor vi børn
stadig gemte alle æbleskrog

KBH Text Regular

Min morfar havde også tømt
sit handskerum, hvor vi børn
stadig gemte alle æbleskrog

Gill Sans Light

Min morfar havde også tømt
sit handskerum, hvor vi børn
stadig gemte alle æbleskrog

Figure 43: KBH Display Regular, KBH Text Regular, and Gill Sans Light, tested by Beier and Oderkerk (2019b). Attribution: Figure 1 from the journal article: Beier, S. and Oderkerk, C.A.T., 2019. The Effect of Age and Font on Reading Ability. *Visible Language,* 53(3), pp.50-68, Copyright (2019), with permission from Visible Language.

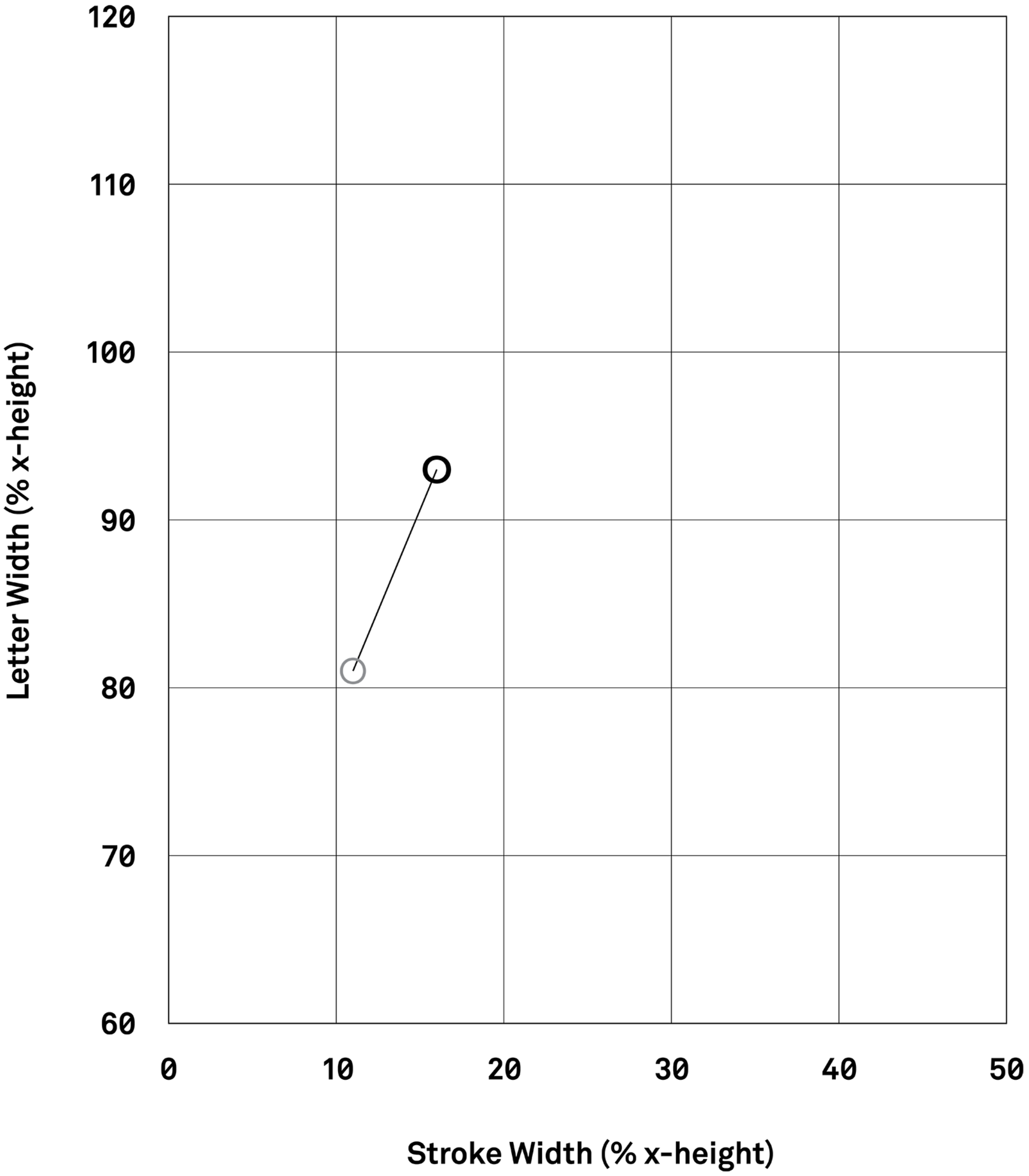


Figure 44: Gill Sans Light is found to be less legible (reading acuity, CPS) than KBH Text Regular for older participants (Beier & Oderkerk, 2019b). KBH Text Regular is represented by an ‘o’ with similar proportions (section 4.2.3).

4.4 DISCUSSION

4.4.1 GENERAL PATTERNS

This review contributes a consolidation of scientific knowledge on the influence of stroke width on legibility in the context of low vision. A visualisation of all typefaces included in the review (Figure 45) illustrates that most of the typefaces with higher legibility (black) have both an increased stroke width and letter width, as compared to the typefaces with lower legibility (grey). Visualising only the relationships between typefaces tested within studies (i.e. lines visualised without ‘o’s) (Figure 46) shows that most of the lines are angled upward and to the right. This again illustrates that typefaces with higher legibility have both increased stroke width and letter width as compared to the typefaces they are tested against. One exception is Memphis Light compared to Memphis Medium (Luckiesh & Moss, 1940) (Figure 28), represented by the only line angled downward to the right (see Figure 46). This suggests that stroke width can improve legibility, even when letter width is not increased. Ovink Ultra Black (not visualised) is also an exception, illustrating that increased stroke width and letter width can also decrease legibility, albeit at an “extreme” weight (Beier & Oderkerk, 2019a, p.4).

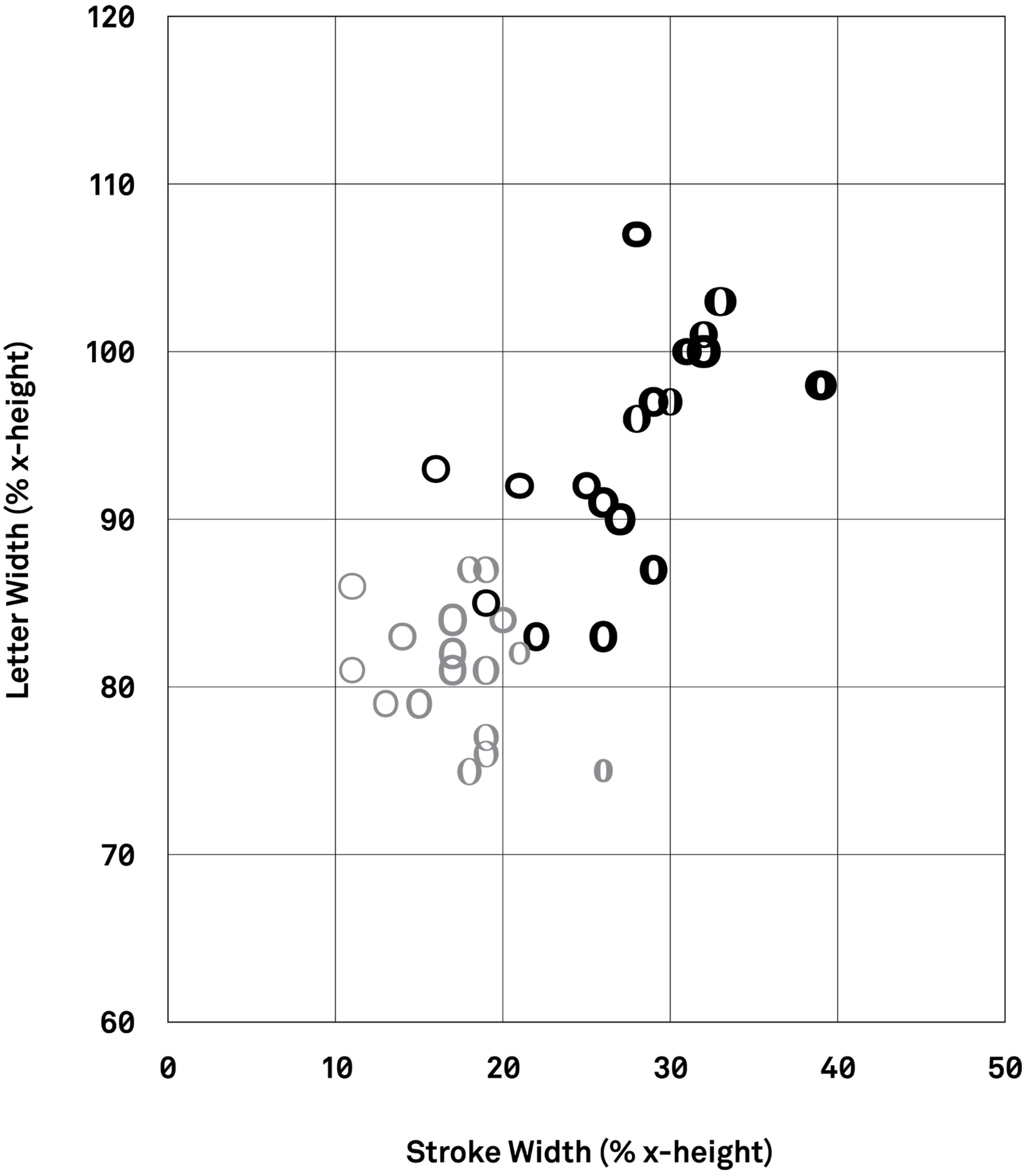


Figure 45: Visualisation of all typefaces found to have lower (grey) and higher (black) legibility included in the review. List of studies: Roethlein, 1912; Luckiesh & Moss, 1940; Shaw, 1969; Smither & Braun, 1994; Mansfield et al., 1996; Sheedy et al., 2005; Beveratou, 2016; Xiong et al., 2018; Beier & Oderkerk, 2019a; Beier & Oderkerk, 2019b.

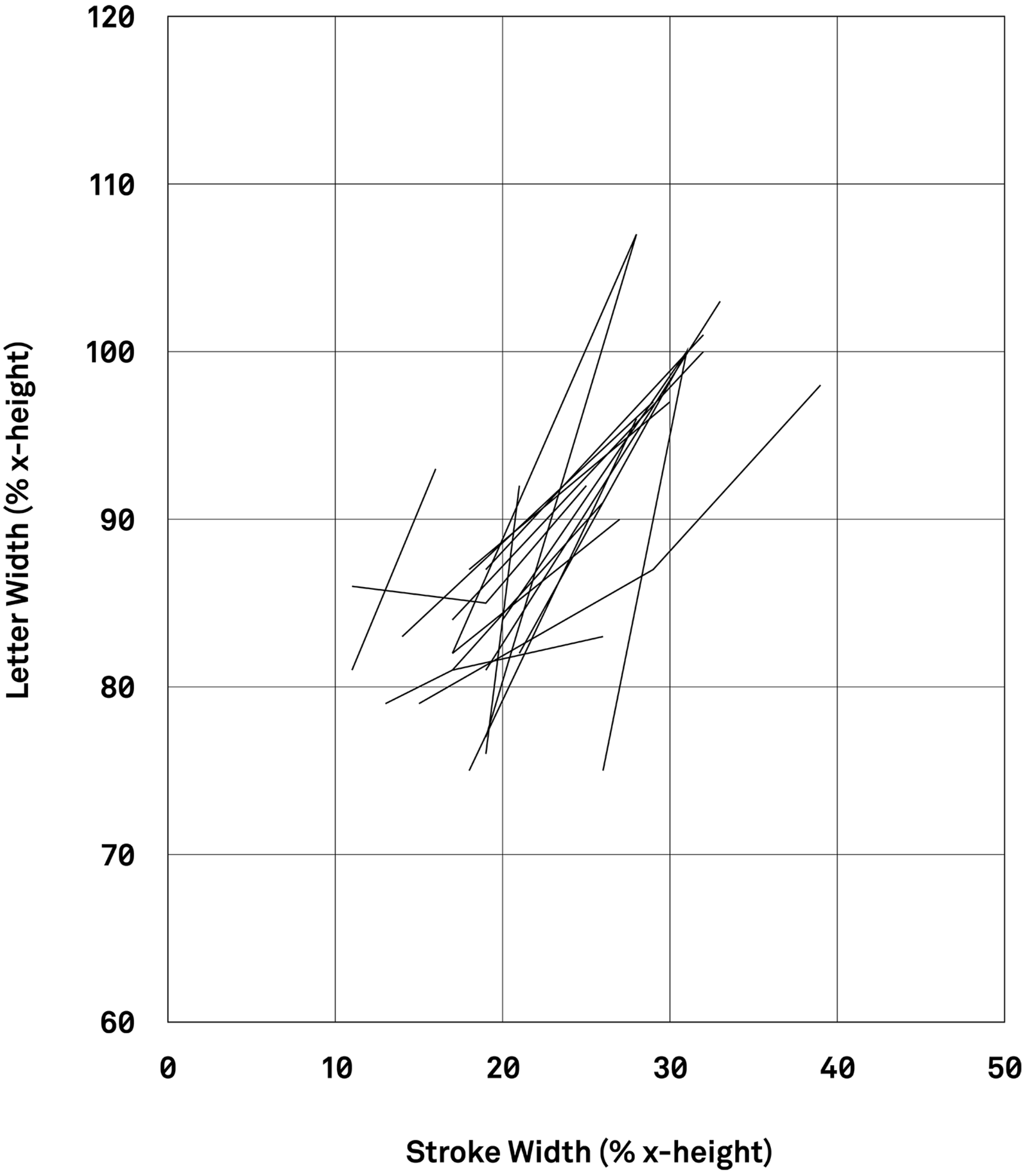
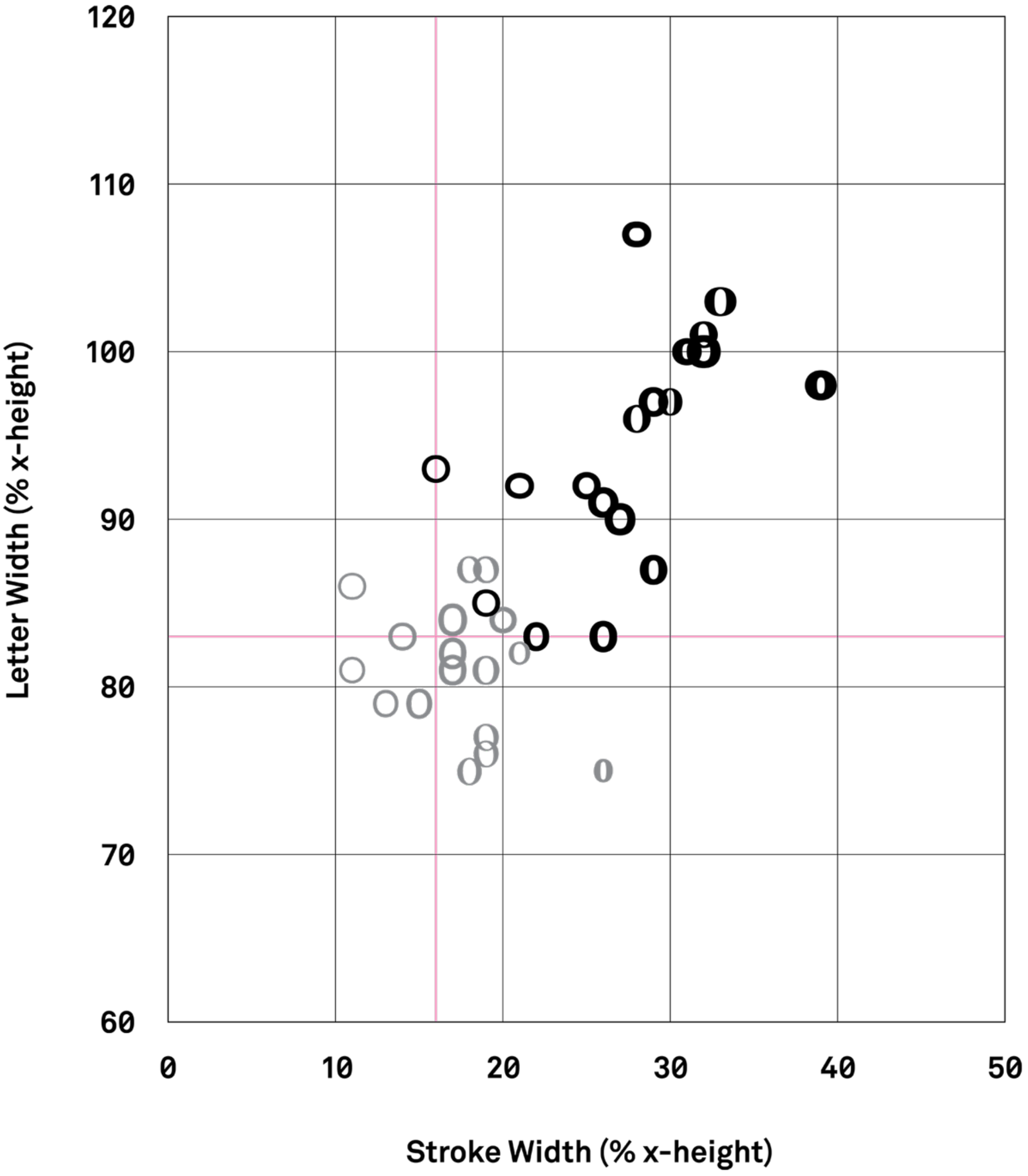


Figure 46: Visualisation of all relationships between typefaces found to have lower and higher legibility included in the review. List of studies: Roethlein, 1912; Luckiesh & Moss, 1940; Shaw, 1969; Smither & Braun, 1994; Mansfield et al., 1996; Sheedy et al., 2005; Beveratou, 2016; Xiong et al., 2018; Beier & Oderkerk, 2019a; Beier & Oderkerk, 2019b.

4.4.2 TYPEFACE STROKE WIDTH AND LOW VISION LEGIBILITY

Chapter 2 (section 2.4.3.3) illustrates that inclusive print guidelines recommend a range of weights from regular to bold (e.g. RNIB & ISTD, 2007). Based on Chapter 4's analysis of scientific studies, the question of what stroke widths to recommend for low vision reading can be addressed. In order to do so, visualisations are created which include two pink lines intersecting at right angles, dividing the data into four quadrants (Figures 47-50). The vertical pink line marks the lowest stroke width value among the typefaces found to have higher legibility (i.e. black 'o's). Similarly, the horizontal pink line marks the lowest letter width value among the typefaces found to have higher legibility. As such, the upper right quadrant contains all typefaces found to have higher legibility (see Figure 47). In this way, the proportions of typefaces with higher legibility are delineated. Specifically, the pink lines indicate the stroke width and letter width values above which we might recommend for low vision reading. Note that while letter width values are demarcated, the discussion focuses on stroke width because it is the purpose of the review. The following analyses show that all ten studies contribute to knowledge on the relationship between stroke width, letter width, and legibility, however they are not all useful in addressing the practical question of stroke width recommendations.

Based on all of the typefaces included in the review (Figure 47), a minimum stroke width of 16% appears advisable, determined by KBH Text Regular (Beier & Oderkerk, 2019b) (Figure 44). However, this study only tests a regular and a light weight typeface, and therefore cannot inform whether typefaces above regular weights (i.e. standard text setting) might improve legibility. Therefore, this study is removed from further analyses.



<u>Figure 47</u>: The upper right quadrant contains typefaces found to have higher legibility (black ‘o’s), based on all studies included in the review. List of studies: Roethlein, 1912; Luckiesh & Moss, 1940; Shaw, 1969; Smither & Braun, 1994; Mansfield et al., 1996; Sheedy et al., 2005; Beveratou, 2016; Xiong et al., 2018; Beier & Oderkerk, 2019a; Beier & Oderkerk, 2019b. KBH Text Regular has the lowest stroke width value among the typefaces found to have higher legibility (Beier & Oderkerk, 2019b) (see Figure 44).

The study by Luckiesh and Moss (1940) is also removed from further analyses. This study only finds a substantial difference in legibility between a light (11% stroke width) and a medium weight (19% stroke width) typeface (Figure 28), in contrast to the other eight studies which find legibility benefits at much higher stroke widths. This suggests that Memphis Medium is an underestimate of optimal stroke width, potentially due to the use of visibility as a measure of legibility. While the Luckiesh-Moss visibility meter can be used to determine visibilities of two objects on a relative scale, it “cannot be used to determine the optimal font size or type” (Subbaram, 2004, p.10). As explained by Tinker (1963, p.10), it is unclear “what scale value corresponds to optimal legibility”.

Figure 48 visualises typefaces from eight studies, with the removal of two studies from the dataset (Luckiesh & Moss, 1940; Beier & Oderkerk, 2019b). Note that the upper right quadrant no longer contains any of the typefaces found to have lower legibility (i.e. grey ‘o’s). This indicates that the dataset based on eight studies offers a better understanding of typeface proportions associated with improved reading performance. Based on this visualisation, a minimum stroke width of 21% appears advisable, determined by Courier Bold (Mansfield et al., 1996) (Figure 33). However, based on the analyses of this bold monospaced typeface (section 4.3.6), its higher legibility is more likely due to increased letter width and letter spacing, rather than stroke width. Therefore, while Courier Bold is found to improve reading performance, it cannot be used to inform stroke width recommendations for proportionally spaced typefaces. Similarly, Maxular Rx Bold is not appropriate for this analysis, as its performance benefit is more likely due to notably larger letter width and letter spacing versus increased stroke width. The studies by Mansfield et al. (1996) and Xiong et al. (2018) are removed from further analyses.

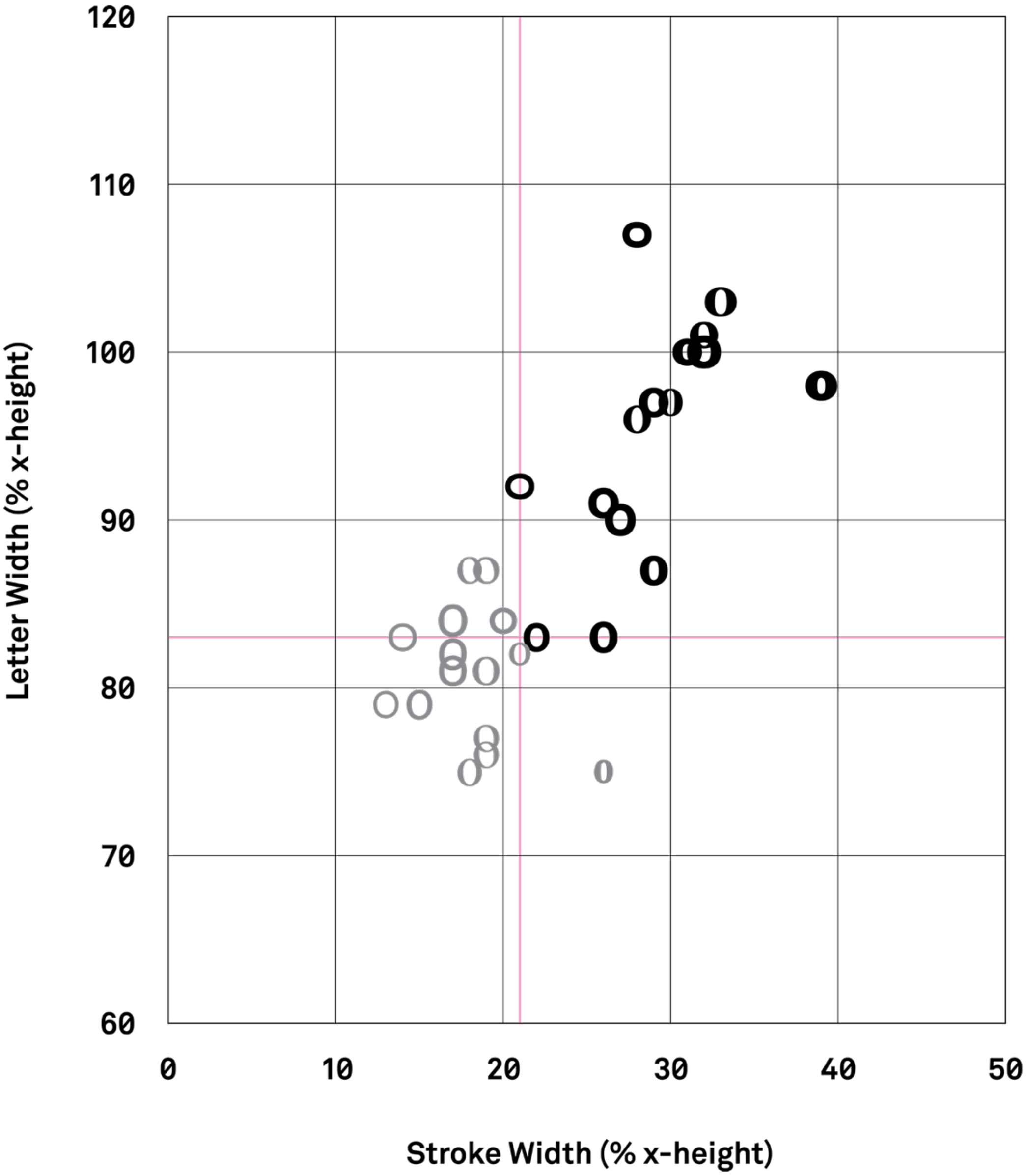


Figure 48: The upper right quadrant contains typefaces found to have higher legibility, based on eight studies included in the review. List of studies: Roethlein, 1912; Shaw, 1969; Smither & Braun, 1994; Mansfield et al., 1996; Sheedy et al., 2005; Beveratou, 2016; Xiong et al., 2018; Beier & Oderkerk, 2019a. Studies testing light weight typefaces are removed from the dataset (Luckiesh & Moss, 1940; Beier & Oderkerk, 2019b). Courier Bold has the lowest stroke width value among the typefaces found to have higher legibility (Mansfield et al., 1996) (see Figure 33).

Figure 49 visualises typefaces from six studies, with the removal of four studies from the dataset (Luckiesh & Moss, 1940; Mansfield et al., 1996; Xiong et al., 2018; Beier & Oderkerk, 2019b). Based on this visualisation, a minimum stroke width of 22% is advisable, determined by Franklin Gothic Medium (Sheedy et al., 2005) (Figure 36). As discussed in section 4.3.7, Sheedy et al. (2005) find that Franklin Gothic Book is less legible (visual acuity) than Franklin Gothic Medium. Based on my calculations undertaken in Chapter 6 (Table 18, section 6.3.2.1), Franklin Gothic Book's stroke width (15%) falls into the lower end of the range of regular weight typefaces. Recommending a 22% stroke width minimum is reasonable, as this would potentially result in a legibility benefit over regular typefaces on the lower end of the stroke width range. Figure 50 allows for a comparison of Figures 47, 48, and 49 (i.e. differing minimum stroke width recommendations).

While 22% stroke width represents a minimum recommendation for low vision reading, the point at which stroke width results in a legibility decline is unclear. Based on the data, legibility benefits are reported with typefaces having quite large stroke width values (Figure 49). While Franklin Gothic Heavy (39% stroke width)—the typeface with the largest stroke width in Figure 49—improves legibility for lowercase text, legibility suffers with uppercase (Subbaram, 2004). As text is generally mixed case (i.e. upper and lowercase), Franklin Gothic Demi (29%) represents the highest stroke width that can be recommended based on Sheedy et al.'s (2005) investigation of Franklin Gothic. Beier and Oderkerk (2019a) also find the highest stroke width resulting in increased legibility is 29%. However, as can be seen from the visualisation in Figure 49, there are typefaces above 29% stroke width that are found to improve legibility.

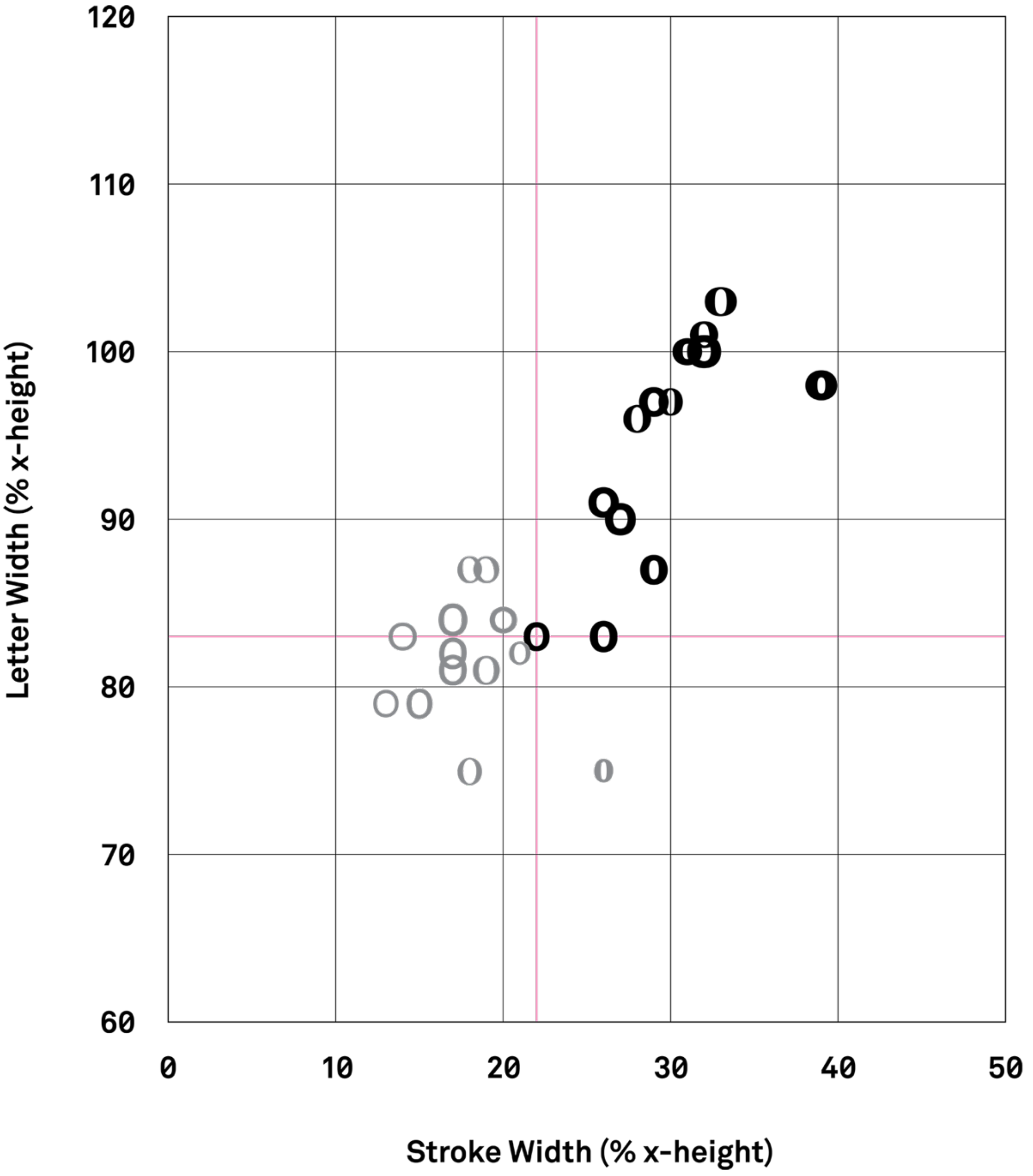
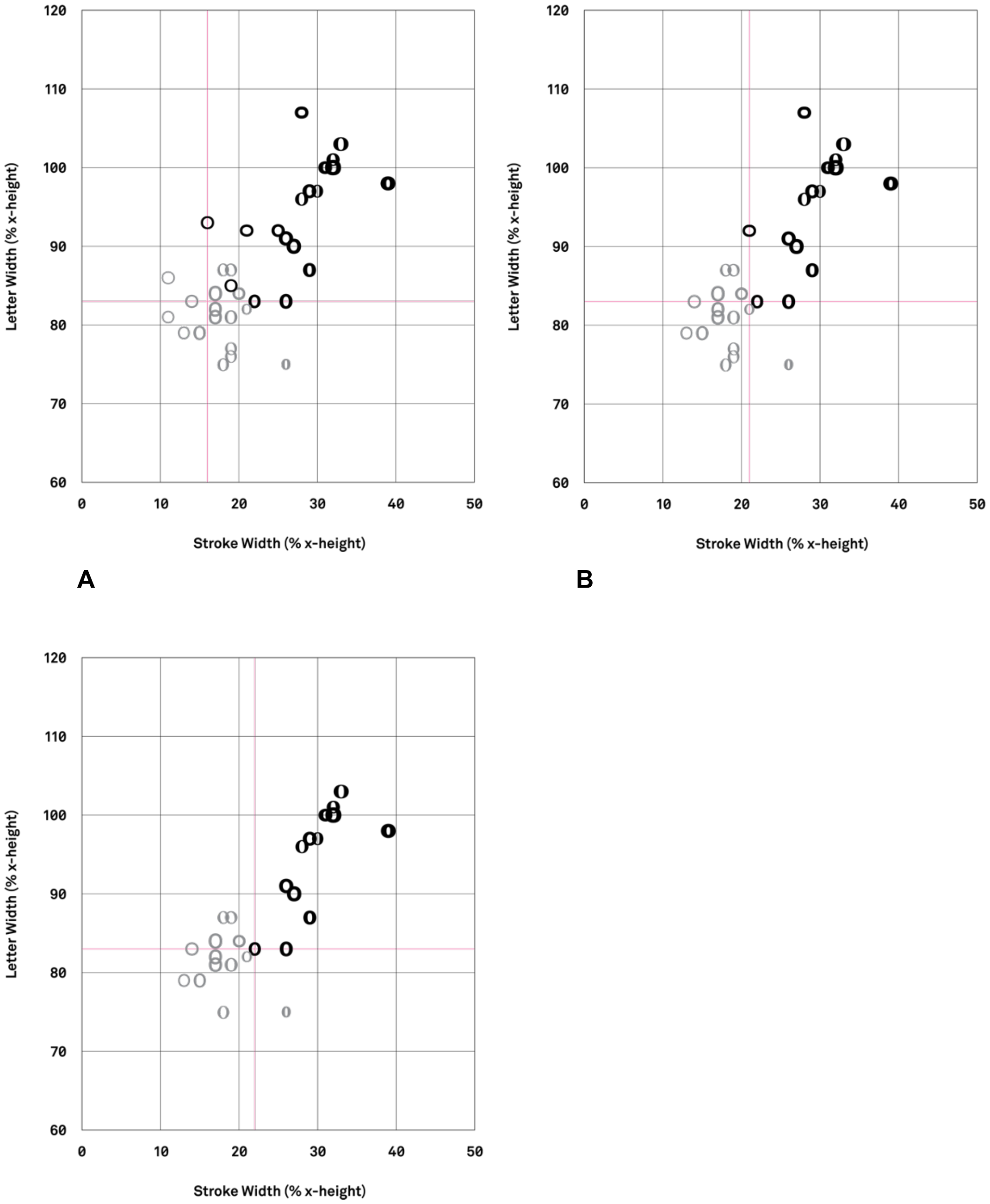


Figure 49: The upper right quadrant contains typefaces found to have higher legibility, based on six studies included in the review. List of studies: Roethlein, 1912; Shaw, 1969; Smither & Braun, 1994; Sheedy et al., 2005; Beveratou, 2016; Beier & Oderkerk, 2019a. Studies testing light weight typefaces (Luckiesh & Moss, 1940; Beier & Oderkerk, 2019b) and typefaces with increased letter width and letter spacing (Mansfield et al, 1996; Xiong et al., 2018) are removed from the dataset. Franklin Gothic Medium has the lowest stroke width value among the typefaces found to have higher legibility (Sheedy et al., 2005) (see Figure 36).



A

B

C

Figure 50: Figures 47, 48, and 49 compared. The upper right quadrants contain typefaces found to have higher legibility. Typefaces are visualised from (A) all studies, (B) studies testing light weight typefaces removed (Luckiesh & Moss, 1940; Beier & Oderkerk, 2019b), and (C) studies testing typefaces with increased letter width and letter spacing also removed (Mansfield et al, 1996; Xiong et al., 2018).

Shaw (1969) and Sheedy et al. (2005) offer insight into the potential upper range of stroke widths. Based on my calculations, Shaw (1969) finds legibility benefits at quite high stroke widths; 31% (Gill Sans Bold) and 32% (Plantin Bold), as compared to typefaces with stroke widths at the higher end of regular weights; 20% (Gill Sans Regular) and 19% (Plantin Regular) (Figure 30). Sheedy et al. (2005) also find legibility benefits at high stroke widths; 32% (Verdana Bold) and 33% (Georgia Bold), compared to stroke widths at 17% (Verdana Regular) and 19% (Georgia Regular) (Figure 35). Based on this data, stroke widths of up to 33% can be recommended.

While stroke widths above 30% may improve legibility, it is not known whether this benefit could be achieved with smaller stroke widths, or whether larger benefits could be achieved with even higher stroke widths. What is the optimal typeface stroke width for low vision adults? This analysis indicates that stroke widths ranging from 22-33% improve legibility as compared to regular typefaces, however where the optimum lies remains unclear. Future legibility research would benefit from testing stroke widths of both intermediate (i.e. between 22-29%) and higher (i.e. above 30%) values, in order to determine an optimum for low vision readers. A stroke width of 33% (e.g. Georgia Bold) is a particularly interesting notion for an optimum, as the stroke widths and counters are roughly equivalent in size (Figure 35).

4.5 SUMMARY

This chapter consolidates scientific knowledge on typeface stroke width in the context of low vision readers. Though a quantitative analysis, the stroke width and letter width of typefaces experimentally found to have higher and lower legibility are visualised. The analysis suggests a 22% stroke width minimum for low vision readers, based on the study by Sheedy et al. (2005). The analysis further indicates that stroke widths up to 33% can improve legibility compared to regular weight typefaces, also based on the study by Sheedy et al. (2005). Regarding the optimal stroke width for low vision readers, the analysis suggests that future legibility research tests stroke widths of both intermediate (i.e. between 22-29%) and higher (i.e. above 30%) values, in order to determine this.

Chapter 4 employs a quantitative analysis of typefaces to consolidate scientific knowledge. This knowledge informs the design of a laboratory typeface, presented in Chapter 5 which follows.

CHAPTER 5. DESIGN OF A LABORATORY TYPEFACE

5.1. INTRODUCTION

This chapter describes the design of a laboratory typeface. The typeface is capable of the controlled investigation of stroke width while reflecting design practice in its construction. This investigation addresses the methodological issues of legibility research reviewed in Chapter 2 (section 2.3). As discussed, legibility research often tests commercial typefaces which differ across parameters, resulting in uncontrolled investigations. While parametric typefaces are capable of controlled study, unconventional forms can make it difficult to apply experimental results to design practice (e.g. Arditi et al., 1995a). Legibility researchers have long advocated for experimental test material to reflect design practice (Tinker, 1965), and there are examples of interdisciplinary research illustrating that laboratory typefaces can be based on traditional forms without compromising experimental control (Morris et al., 2002; Bessemans, 2016a; Beier & Dyson, 2014).

The parametric typeface presented in this chapter is informed by scientific and design knowledge. Scientific knowledge is consolidated through the analysis undertaken in Chapter 4. As the laboratory typeface was designed in 2009, a subset of these publications (i.e. pre-2009) are reviewed in this chapter, with a focus on sans serifs. This entails analysing the proportions of typefaces tested in these legibility studies through information visualisation. Design knowledge is also generated in this chapter through a design phenomenology study. This entails measuring and visualising the proportions of a selection of text typeface families. This scientific and design knowledge is used to inform the development of a laboratory typeface capable of testing the proportions of typefaces found to improve reading performance (i.e. scientific knowledge), and those reflecting typeface design practice (i.e. design knowledge). This chapter describes the consolidation of scientific knowledge and generation of design knowledge, and how this knowledge is used to inform the development of the laboratory typeface.

Multiple master technology is used to create the laboratory typeface family. Nine typefaces are designed which represent a matrix of three stroke widths and three letter widths. As discussed in Chapters 1-3, bolder typefaces generally have larger letter widths, therefore in order to investigate three stroke widths, three letter widths are also investigated. Through the design of these nine typefaces, the isolated influence of stroke width and letter width on legibility can theoretically be examined.

5.2 METHODS

5.2.1 CONSOLIDATING SCIENTIFIC KNOWLEDGE

Experimental studies conducted prior to 2009 inform the parameters of the laboratory typeface. This consolidation of scientific knowledge includes a subset of the studies reviewed in Chapter 4 and is focused on sans serif typefaces utilised as test material. The focus on sans serif typefaces is based on evidence suggesting that they improve reading performance for people with low vision (e.g. Morris et al., 2002), presented in Chapter 2 (section 2.4.3.2). This subset of studies and typefaces is appropriate to inform the laboratory typeface which is sans serif.

The subset of typefaces analysed is listed in Table 13. These five sans serif typeface families are also included in the generation of design knowledge, presented in the following section. While the laboratory typeface is based on this subset of studies, it is compared to the results of the full review conducted in Chapter 4 in the discussion (section 5.5.1).

Table 13: Scientific studies informing the design of the laboratory typeface, and typefaces tested experimentally within each. See Appendix 2 for parameter data.

Scientific Study	**Typeface Tested**
Shaw (1969)	Gill Sans Roman
	Gill Sans Bold
Smither & Braun (1994)	Helvetica
	Helvetica Bold
Sheedy et al. (2005)	Arial
	Arial Bold
	Verdana
	Verdana Bold
	Franklin Gothic Book
	Franklin Gothic Medium
	Franklin Gothic Demi
	Franklin Gothic Heavy

5.2.2 GENERATING DESIGN KNOWLEDGE

To ensure that the laboratory typeface reflects design knowledge, an investigation into the proportions of text typefaces is undertaken. The investigation focuses on eighteen text typeface families classified as either grotesque (e.g. Akzidenz-Grotesk), neo-grotesque (e.g. Helvetica), or humanist sans serif (e.g. Frutiger). A typeface classified as neo-modern—used to typeset my visualisations—is also included (Akkurat). The text typeface families analysed in this chapter are listed in Table 14, including those also utilised to consolidate scientific knowledge as described in section 5.2.1.

The proportions of twenty-three text typeface families (193 individual typefaces) are measured and visualised in order to inform the design of the laboratory typeface. This analysis is focused on stroke width and letter width, however many typefaces are also studied less formally to inform other aspects of the laboratory typeface. This includes typeface characteristics such as stroke contrast, shape of counters, and widths of other letters in the alphabet. Where necessary, detailed analyses are focused on Helvetica and Helvetica Neue as these typeface families traverse a large range of weights and widths, and represent some of the most widely used sans serifs (Unger, 2005). Helvetica is also recommended for visually impaired audiences (RNIB, 2017).

Table 14: Twenty-three sans serif typefaces investigated to generate design knowledge. The 'x' indicates whether roman, condensed, or extended typefaces are analysed. Asterisks indicate typefaces tested within scientific studies. Helvetica is listed twice because Helvetica* refers to a system font and Helvetica refers to a large typeface family.

Typeface Family	Roman	Condensed	Extended
AG Buch	x	x	x
Akkurat	x	-	-
Akzidenz-Grotesk	x	x	x
Arial*	x	-	-
DIN	x	x	-
Folio	x	x	x
Franklin Gothic*	x	x	-
Frutiger	x	x	-
Gill Sans*	x	-	-
Helvetica*	x	-	-
Helvetica	x	x	x
Helvetica Neue	x	x	x
Lucida Sans	x	-	-
Meta	x	x	-
Monotype Grotesque	x	-	-
Myriad Pro	x	x	-
Neuzeit S	x	-	-
News Gothic	x	x	-
Slate Mono	x	-	-
Trade Gothic	x	x	x
Trebuchet	x	-	-
Univers	x	x	x
Verdana*	x	-	-

5.2.3 VISUALISING TYPEFACE DATA

Typefaces are measured employing the same methods described in Chapter 4 (section 4.2.2). Typefaces analysed to consolidate scientific knowledge are also visualised using the same methods as Chapter 4 (section 4.2.3). While typefaces analysed to generate design knowledge are visualised using similar methods, there are also important differences which are described in the following paragraphs. Visualisation of the laboratory typeface’s parameters is also discussed.

The generation of design knowledge involves the visualisation of 193 individual typefaces. Due to the large number of typefaces analysed, and in part due to rounding data to the nearest percent, some of these typefaces have identical stroke width and letter width values. When visualising this data, whenever two ‘o’s overlap, one is deleted. This occurs in twelve instances. While these typefaces have similar proportions, they differ visually, therefore the overlapping ‘o’s would have obfuscated each letterform. The removal of one ‘o’ is necessary to improve the clarity of the visualisations. As the typefaces are similar in proportion, this choice has little effect on communicating the relationship between typeface parameter values and letterform design.

CMYK colours are once again utilised in the visualisation of design knowledge. The typefaces are initially presented as black (Figures 52a, b, and c), and then colour is employed in subsequent visualisations (Figures 53 and 54) to represent typefaces with different width (e.g. condensed) and weight (e.g. bold) names. Within these more colourful visualisations, a consistent strategy is used. Typeface categories that are most commonly used to set text are presented in black, red represents the main categories of type with increased letter width and stroke width, and blue represents categories with decreased letter width and stroke width. For example, the visualisation of typeface width nomenclature (Figure 53) represents roman typefaces as black (100%K), extended typefaces as red (100%M, 100%Y), and condensed typefaces as blue (100%C). The visualisation of typeface weight nomenclature (Figure 54) represents regular typefaces as black (100%K), bold typefaces as red (100%M, 100%Y), and light typefaces as blue (100%C). Further weight categories are visualised as follows: medium weight typefaces are represented by pink (50%M); extra-bold typefaces are represented by violet (50%C, 50%M). Note that heavier typefaces are layered in Adobe Illustrator above lighter weights, therefore if ‘o’s of different colours overlap, the heavier weight category is visually unobscured. Regarding the visualisation of weight nomenclature, the five categories of weights represent many more weight names. The weights are grouped as follows: Light (Light, Extra-Light, Ultra-Light, Thin); Regular (Book, Normal, Regular, Roman); Medium (Demi, Demi-Bold, Medium, Semi Bold); Bold; Extra-Bold (Black, Book Heavy, Extra-Black, Extra-Bold, Heavy, Super, Ultra-Black).

The parameters of the laboratory typeface are also visualised in this chapter, illustrating how they relate to design knowledge (e.g. Figure 57) and scientific knowledge (e.g. Figure 60). A black rectangle delineates the maximum and minimum weights and widths of the typeface (e.g. Figure 57). Nine black circles are also used to represent the parameters of each of the nine typefaces (e.g. Figure 56).

5.2.4 PARAMETRIC TYPEFACE DESIGN

The laboratory typeface is designed using multiple master technology. Multiple master typefaces contain several styles called masters in one file (FontLab, 2006). From this file one can select to use any of the master typefaces, as well as any intermediate style created by interpolation of the masters. Multiple master technology is employed to design the laboratory typeface across two axes of interest: weight and width. This requires designing four master typefaces at the extremes of the weight and width values of interest, and interpolating typefaces with intermediate values. Further details regarding the design of the laboratory typeface are described in sections 5.3.3 and 5.4.

5.3 RESULTS

5.3.1 CONSOLIDATING SCIENTIFIC KNOWLEDGE (1969-2005)

Based on the visualisation (Figure 51) of sans serif typefaces experimentally tested by Shaw (1969), Smither and Braun (1994) and Sheedy et al. (2005), three major insights inform the parameters of the laboratory typeface. First, the typefaces with higher legibility are all above 22% stroke width. Therefore, the laboratory typeface should be designed to test stroke widths above and below this value to investigate the influence of stroke width on legibility. Second, typefaces with stroke widths as high as 31% (Shaw, 1969) and 32% (Sheedy et al., 2005) are found to improve legibility, despite their smaller counters. Therefore, the laboratory typeface should be designed to test relatively high stroke widths, above 30%. Third, typefaces at high stroke widths of 31-32% have letter widths of 100%, meaning that the width of the ‘n’ is as tall as it is wide.

While Subbaram (2004) finds legibility benefits for lowercase Franklin Gothic Heavy (39% stroke width), this is considered too extreme a stroke width for the laboratory typeface. At this stroke width and letter width, the counter is smaller than the strokes (see Figure 51: ‘o’ with highest stroke width value). The visualisation of design knowledge (Figure 54) discussed in the following section also illustrates that most bold typefaces are below 39% stroke width. This highlights one of the design goals, which is to reflect the proportions of commonly used typefaces versus extreme proportions.

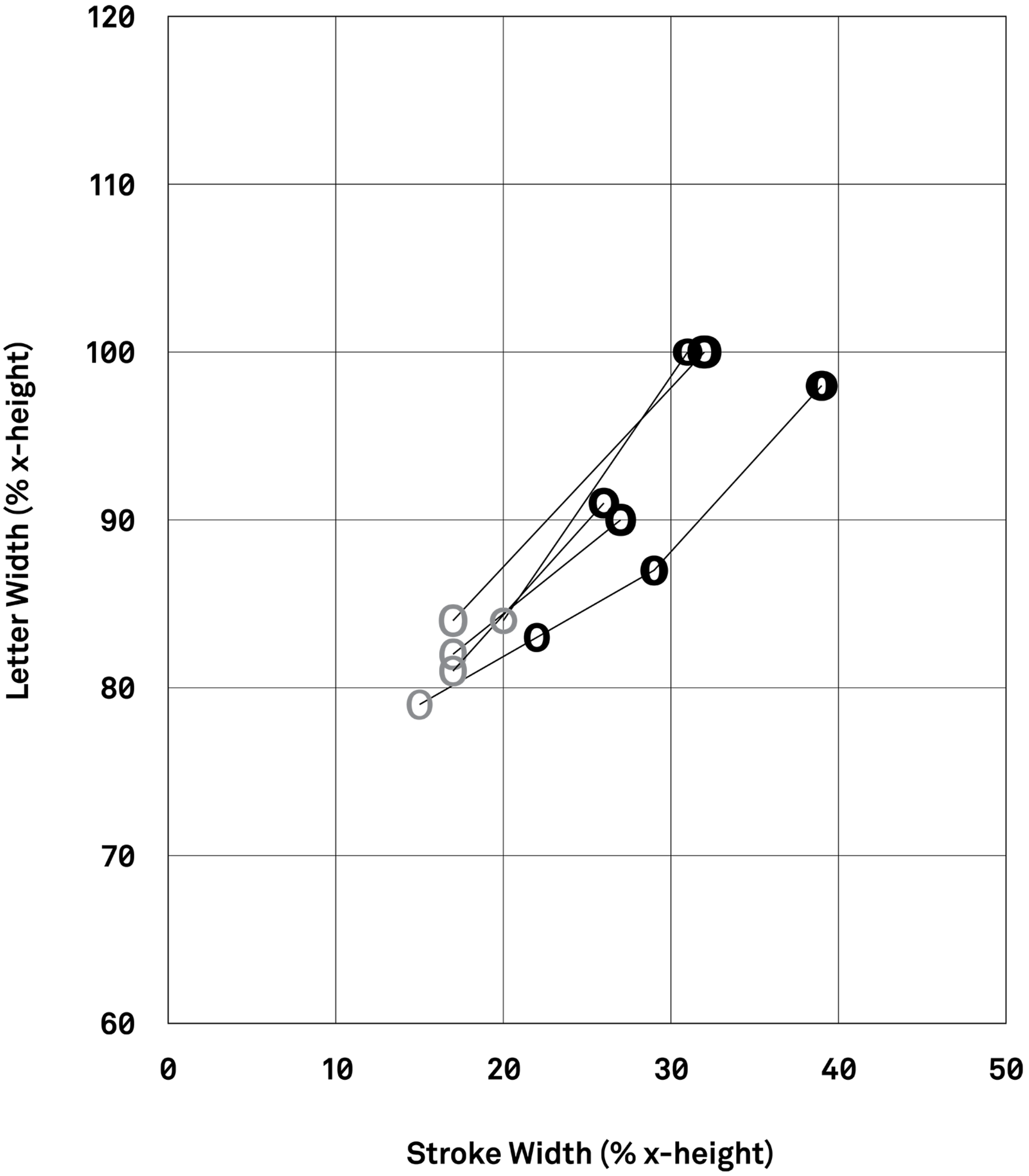


Figure 51: Subset of scientific studies informing the laboratory typeface. Chosen studies pre-date the design of the laboratory typeface (created in 2009) and are focused on sans serif test material. List of studies: Shaw, 1969; Smither & Braun, 1994; Sheedy et al., 2005.

5.3.2 GENERATING DESIGN KNOWLEDGE

The stroke width and letter width of text typeface families are investigated in order to inform the proportions of a laboratory typeface that reflects design knowledge. This section presents the main insights from this generation of design knowledge. The details regarding how this knowledge informs the parameters of the laboratory typeface are described in section 5.3.3.

Figure 52a visualises the stroke width and letter width of twenty-three sans serif typeface families (Table 14). This graph is larger than those presented in previous figures, due to the y-axis (letter width) having a larger scale. This is necessary to visualise all three width categories: roman, extended, and condensed. The diversity of typeface proportions visualised in Figure 52a is large, ranging from light condensed typefaces with low stroke width and letter width values (i.e. bottom left of the graph), to heavy extended typefaces with high values in both parameters (i.e. top right of the graph). This diversity is communicated through the position of the data points and the form of the ‘o’s which are plotted. Figure 52b visualises the data without the axes, grid lines, and labels. The visualisation remains understandable due to the plotting of ‘o’s. This simplified visualisation offers an opportunity for viewers to appreciate the letterforms without other elements. Figure 52c offers a detailed view of 52b; another opportunity to appreciate the diversity of the letterforms.

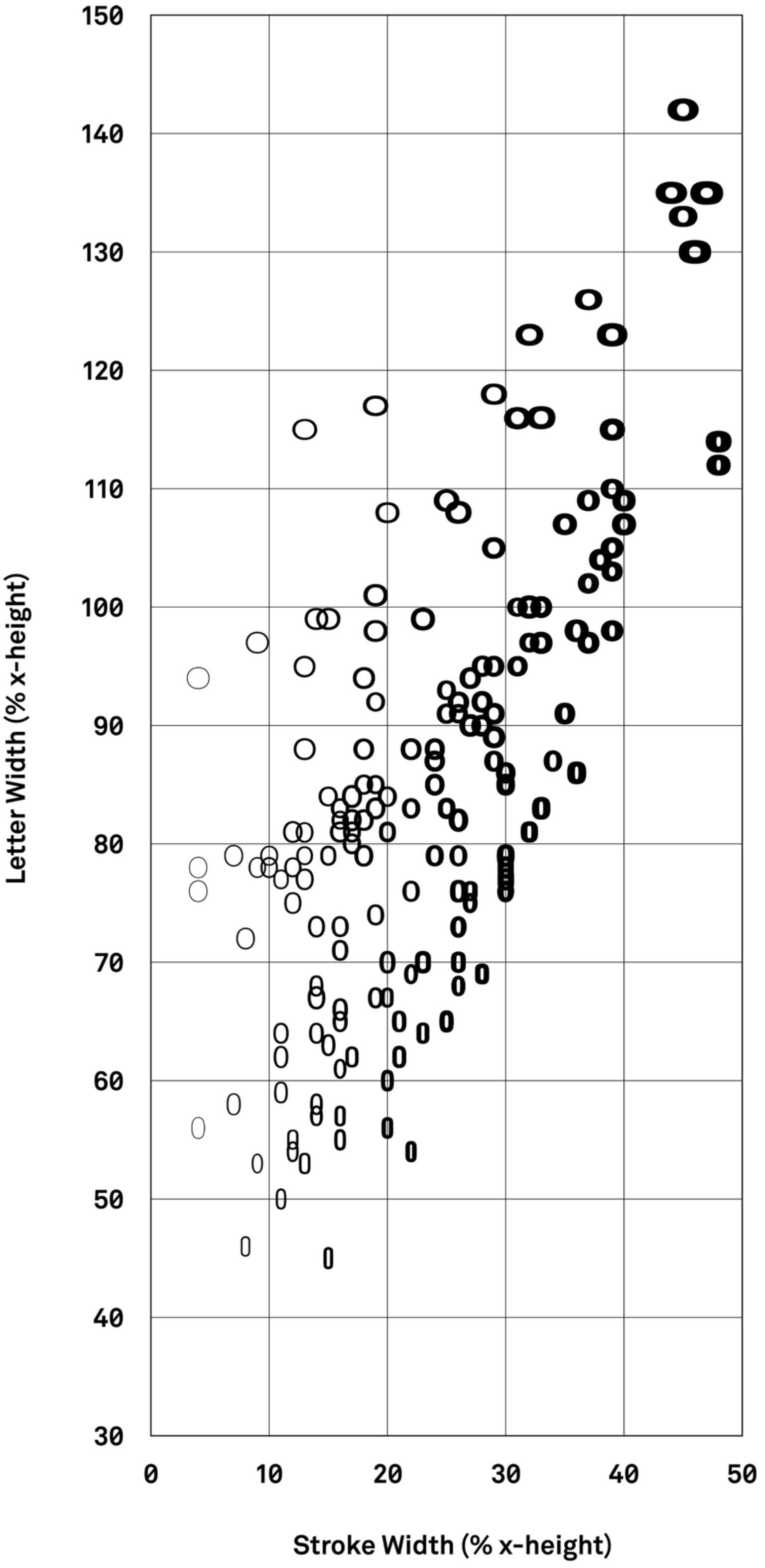


Figure 52a: Stroke width and letter width values of twenty-three sans serif text typeface families. See Table 14 for typefaces analysed.

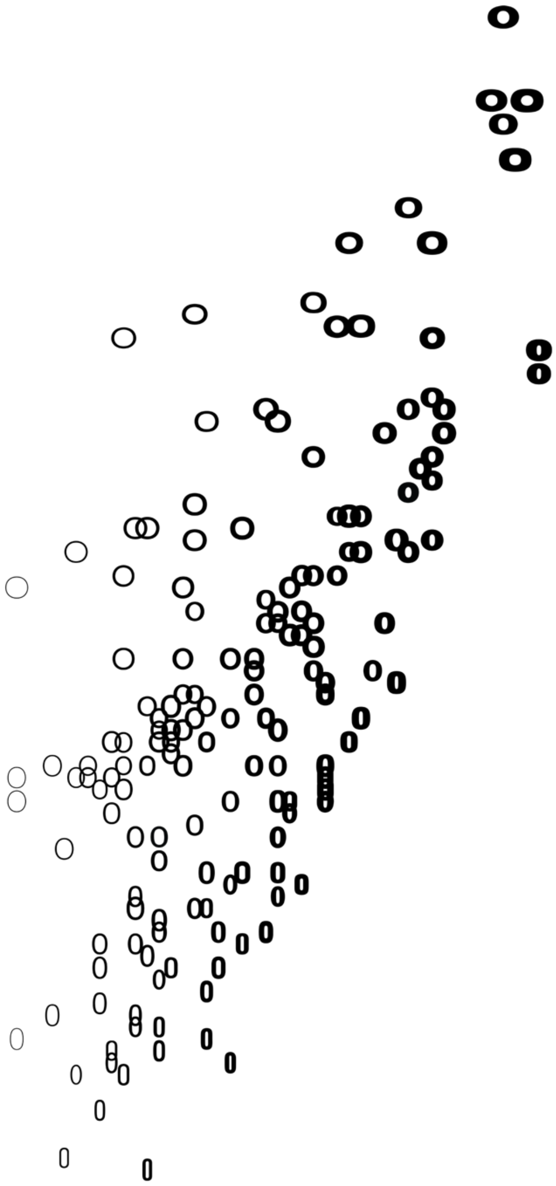


Figure 52b: Stroke width and letter width values of twenty-three sans serif text typeface families with axes, grid lines, and labels removed. See Table 14 for typefaces analysed.

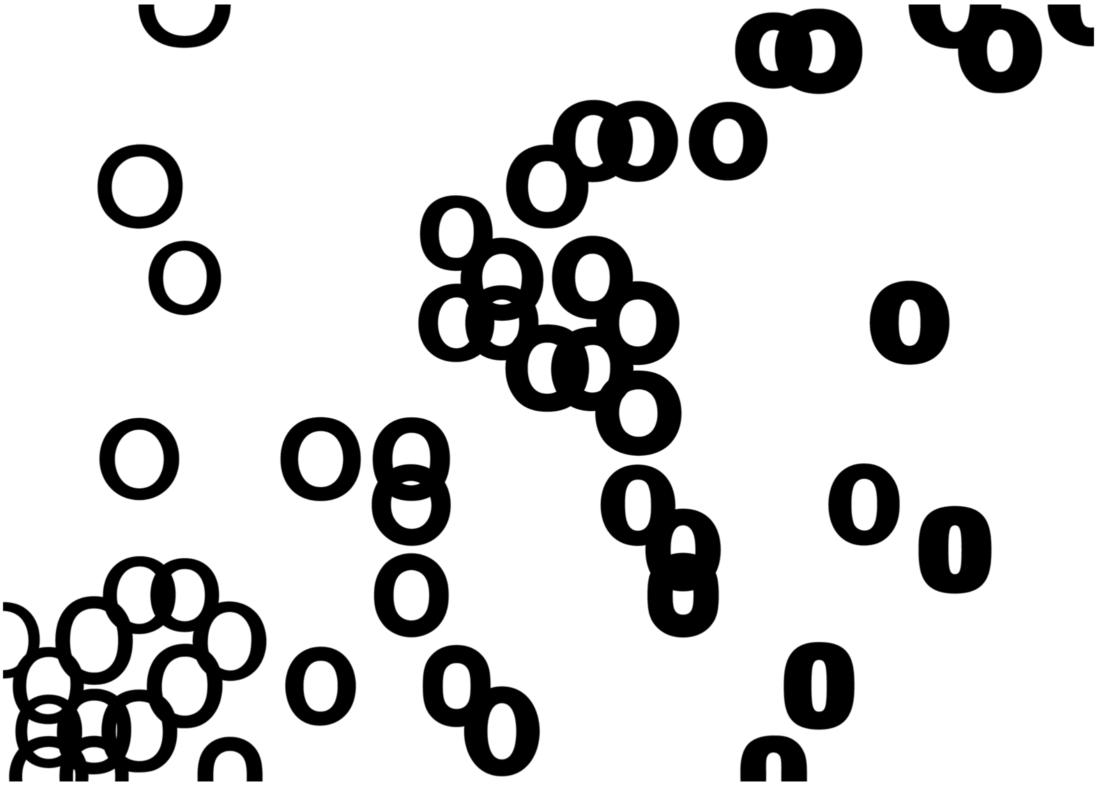


Figure 52c: Detail of Figure 52b.

Figure 53 visualises typeface width nomenclature. As discussed in the methods section 5.2.3, roman typefaces are represented by black, extended by red, and condensed by blue. The visualisation illustrates that in general, typefaces with larger stroke widths tend to have larger letter widths.

Figure 54 visualises typeface weight nomenclature. As discussed in the methods section 5.2.3, light weight typefaces are represented by blue, regular by black, medium by pink, bold by red, and extra-bold by violet. The range of stroke width values of regular (black) and bold (red) typefaces illustrates the lack of standardisation in typeface weight nomenclature, which is considered again in Chapter 6 (section 6.3.2.1). Note that while most of the visualised weight categories in Figure 54 represent a few weight names (see section 5.2.3), the regular category (Book, Normal, Regular, Roman) appropriately represents typefaces for the setting of continuous text, and the bold category only represents typefaces with the name “bold”.

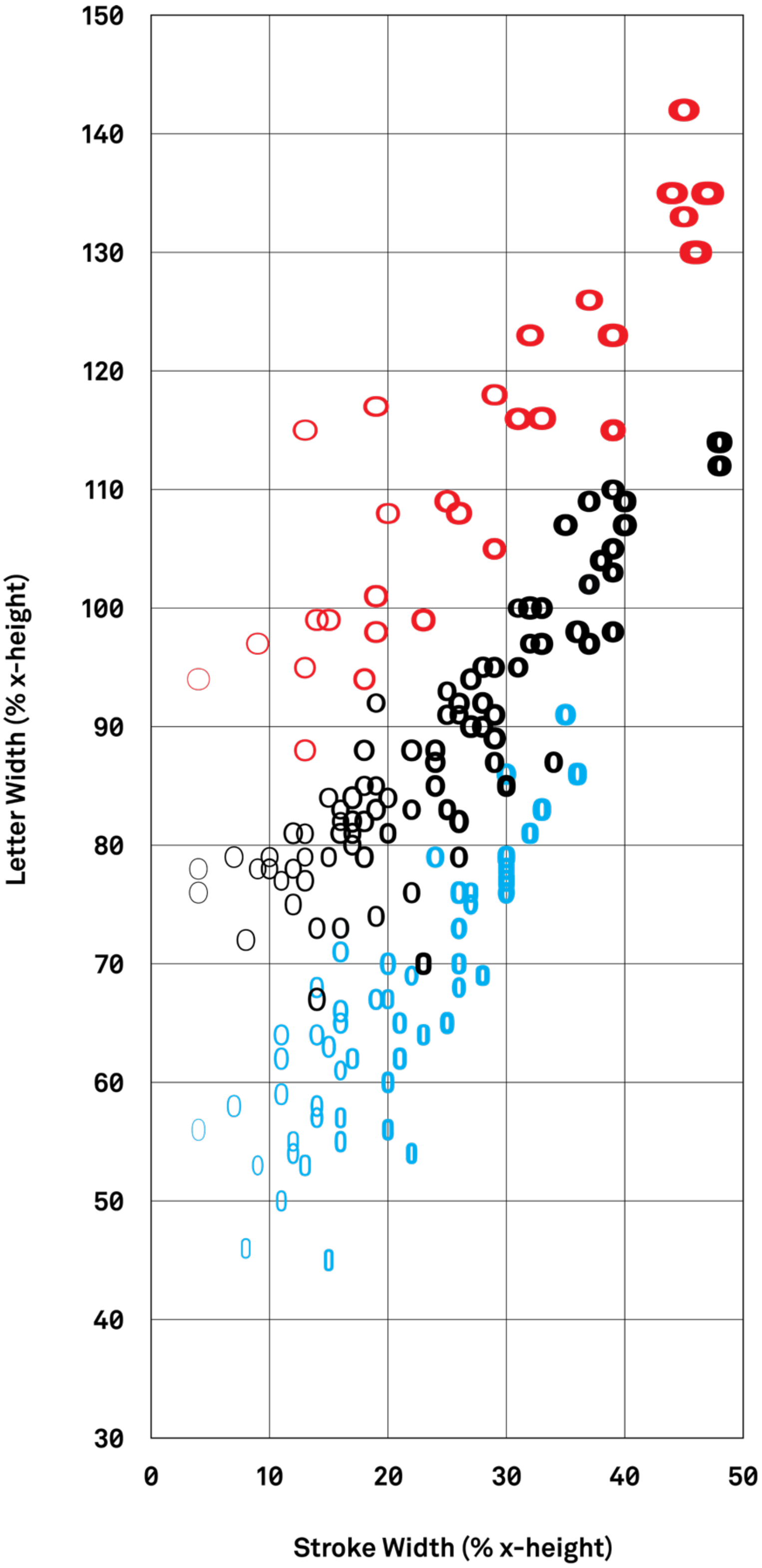


Figure 53: Typeface width nomenclature. Typefaces colour-coded to represent roman (black), extended (red), and condensed (blue) typefaces.

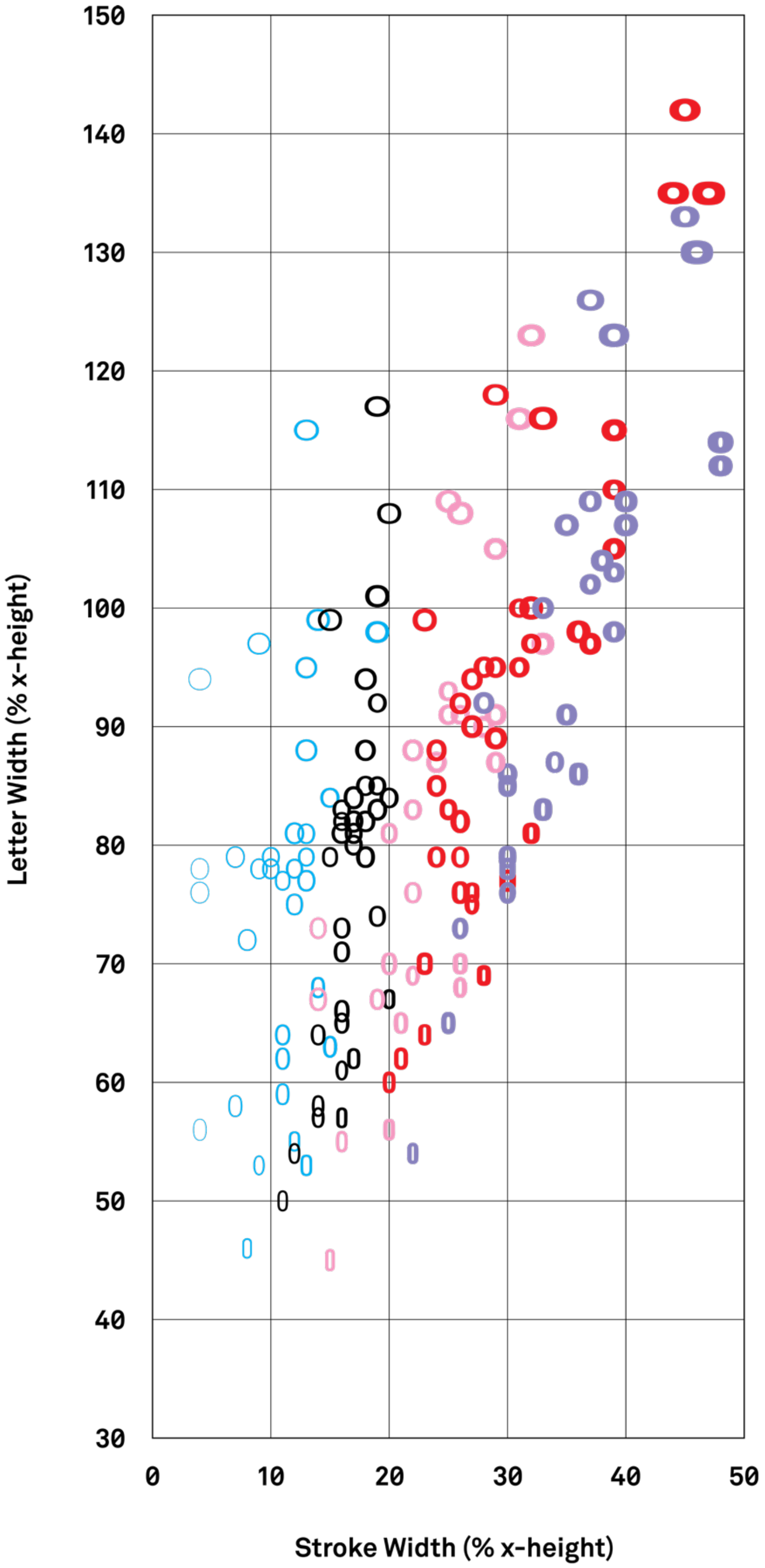


Figure 54: Typeface weight nomenclature. Typefaces colour-coded to represent light (blue), regular (black), medium (pink), bold (red), and extra-bold (violet) typefaces.

5.3.3 LABORATORY TYPEFACE PARAMETERS

Early in my PhD research, I planned to design a laboratory typeface capable of testing three stroke widths by three letter widths; a matrix resulting in a total of nine typefaces. Choosing the exact stroke width and letter width values proved difficult. My earliest iteration proposed to study three stroke widths (10%, 20%, 30%) and three letter widths (80%, 100%, 120%). At this stage I had not done any research into common typeface proportions. This is evident, as there are few typefaces with letter widths as high as 120% (see Figure 53), and a 10% stroke width is quite low (see Figure 54). After I began measuring typeface proportions, I presented a more refined iteration at my first interim exam (May 2008) and ATypI 08 St Petersburg (von Ompteda, 2008), proposing to study three stroke widths (15%, 25%, 35%) and three letter widths (80%, 90%, 100%). While these parameter values were a better reflection of design knowledge, I was only able to make a final decision on the laboratory typeface parameters once I had systematically measured and visualised scientific and design knowledge, as described in sections 5.3.1 and 5.3.2, respectively.

As described in the methods section 5.2.4, the laboratory typeface is designed using multiple master technology. This requires designing four master typefaces at the extremes of the weight and width values of interest, and then interpolating typefaces at intermediate values. Technically an infinite number of typefaces can be interpolated, with extrapolation also possible. The final multiple master typeface is capable of testing three stroke widths (16%, 24%, 32%) and three letter widths (80%, 90%, 100%).

Table 15 lists information about each of the four masters. The masters are defined according to their weight and width, for example wt0wt0 means weight zero width zero. This master has the lowest stroke width (16%) and the lowest letter width (80%), and its proportions correspond to a roman regular typeface. The other masters are listed in Table 15, corresponding approximately to a condensed bold, extended regular, and a roman bold typeface.

Table 15: Master typeface names, meanings, parameter values, and correspondence to commercial typefaces.

Master	Weight and Width	Stroke Width (% x-height)	Letter Width (% x-height)	Correspondence to Commercial Typefaces
wt1wd1	weight one width one	32%	100%	Roman Bold
wt0wd1	weight zero width one	16%	100%	Extended Regular
wt1wd0	weight one width zero	32%	80%	Condensed Bold
wt0wd0	weight zero width zero	16%	80%	Roman Regular

The parameters of the four master typefaces are visualised by black circles in Figure 55. Figure 56 visualises all nine typefaces (i.e. four masters and five interpolated typefaces), and utilises a black rectangle to delineate the maximum and minimum stroke width and letter width values of the laboratory typeface. This rectangle is utilised in the subsequent visualisations to illustrate how the laboratory typeface parameters relate to design knowledge (e.g. Figure 57) and scientific knowledge (e.g. Figure 60).

As the goal is to create a laboratory typeface that reflects design knowledge, it is essential that its parameters reflect the proportions of text typeface families in common use, which are the roman typefaces (versus extended and condensed). The relationship between the laboratory typeface parameters and roman text typefaces is visualised in Figure 57. Figure 57 visualises the laboratory typeface parameters delineating an area that includes most of the regular (black) and bold (red) roman typefaces. It is important that the laboratory typeface is capable of testing regular and bold typeface proportions, as these weights are recommended for readers with low vision (e.g. RNIB & ISTD, 2007). The following paragraphs illustrate the specific ways in which these visualisations inform the design of the laboratory typeface.

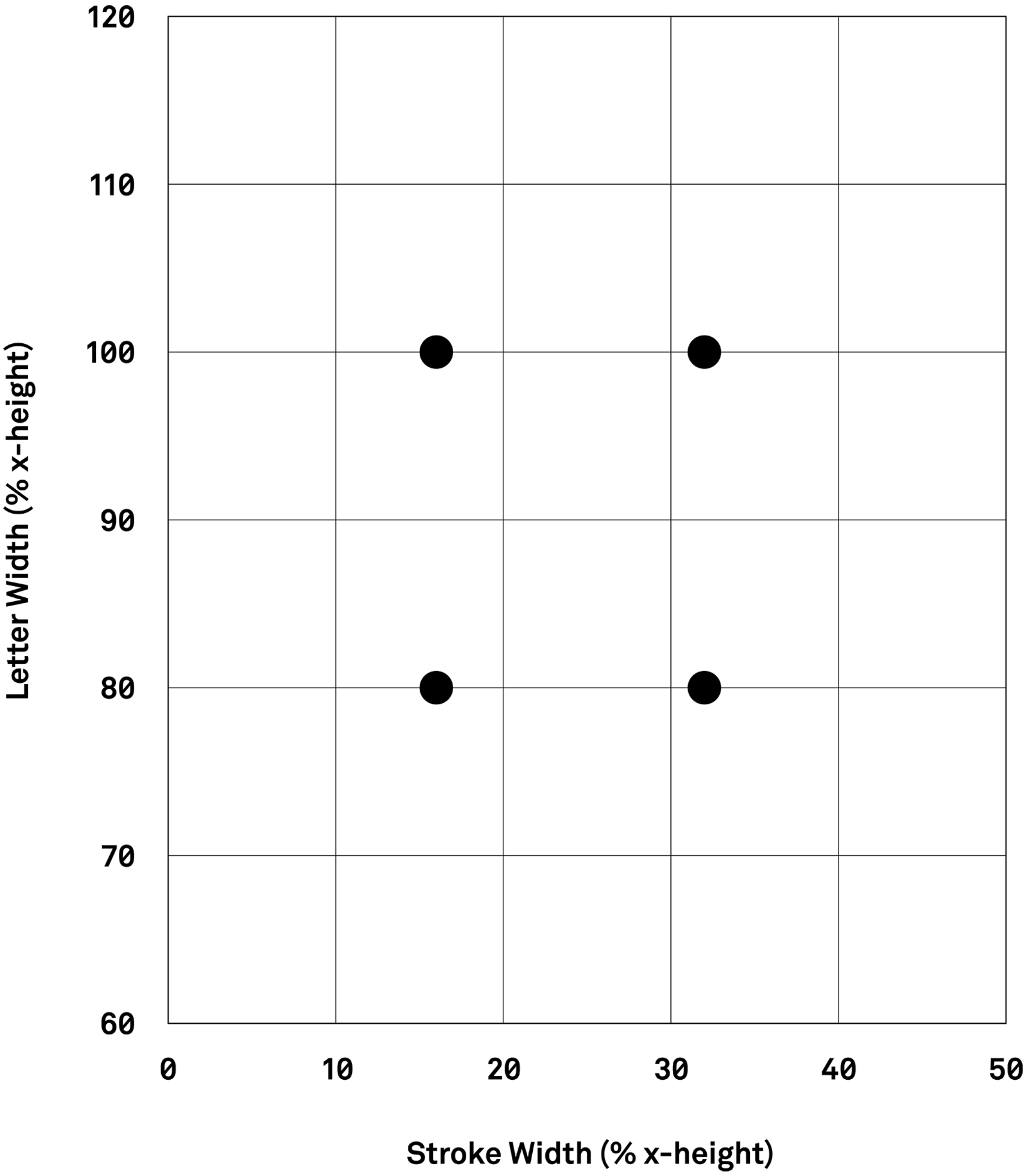


Figure 55: Stroke width and letter width values of the four master typefaces.

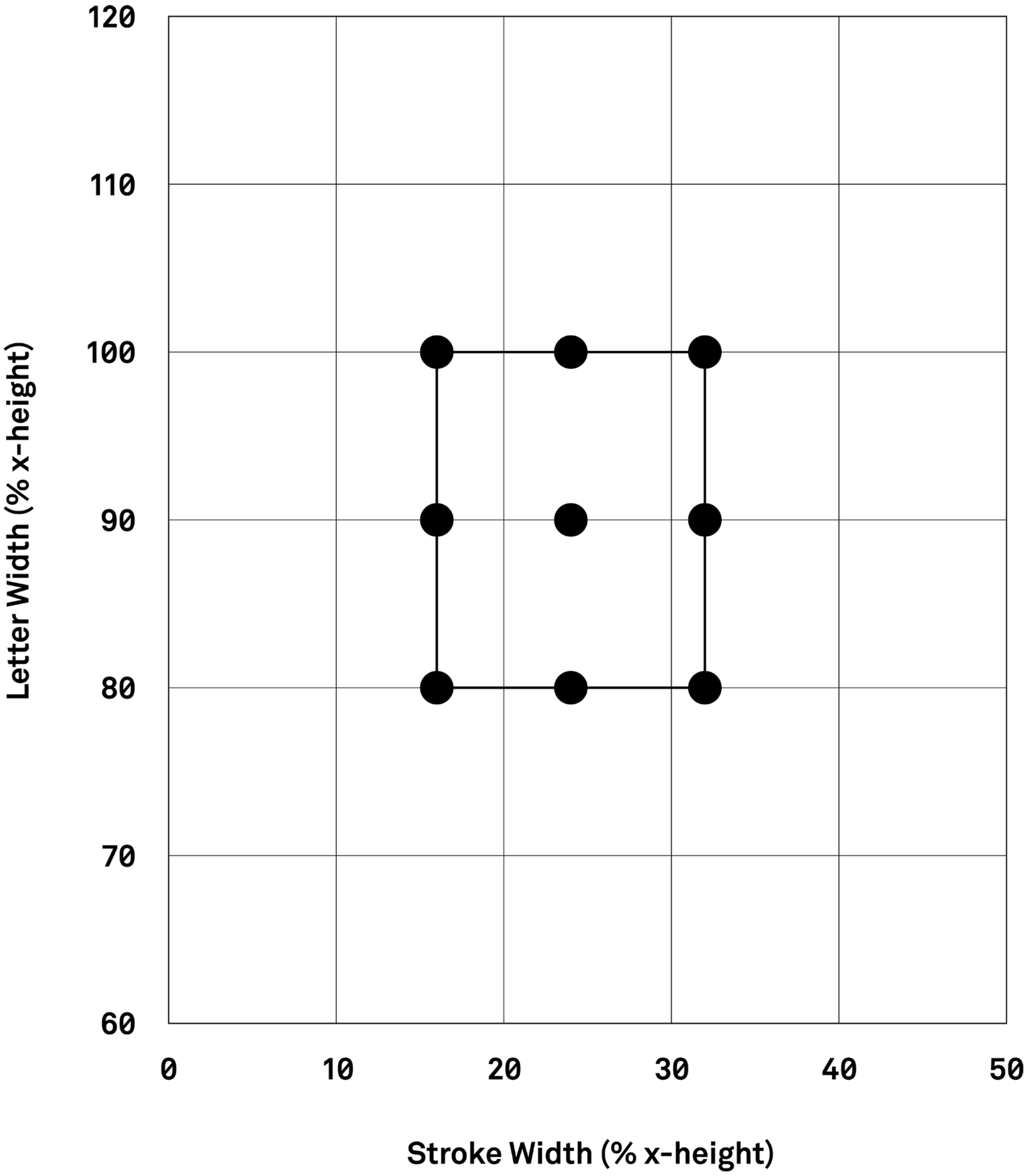


Figure 56: Stroke width and letter width values of the nine laboratory typefaces; four masters and five interpolated typefaces (see Figure 55 for parameter values of four masters). The black rectangle delineates the maximum and minimum stroke width and letter width values of the typeface.

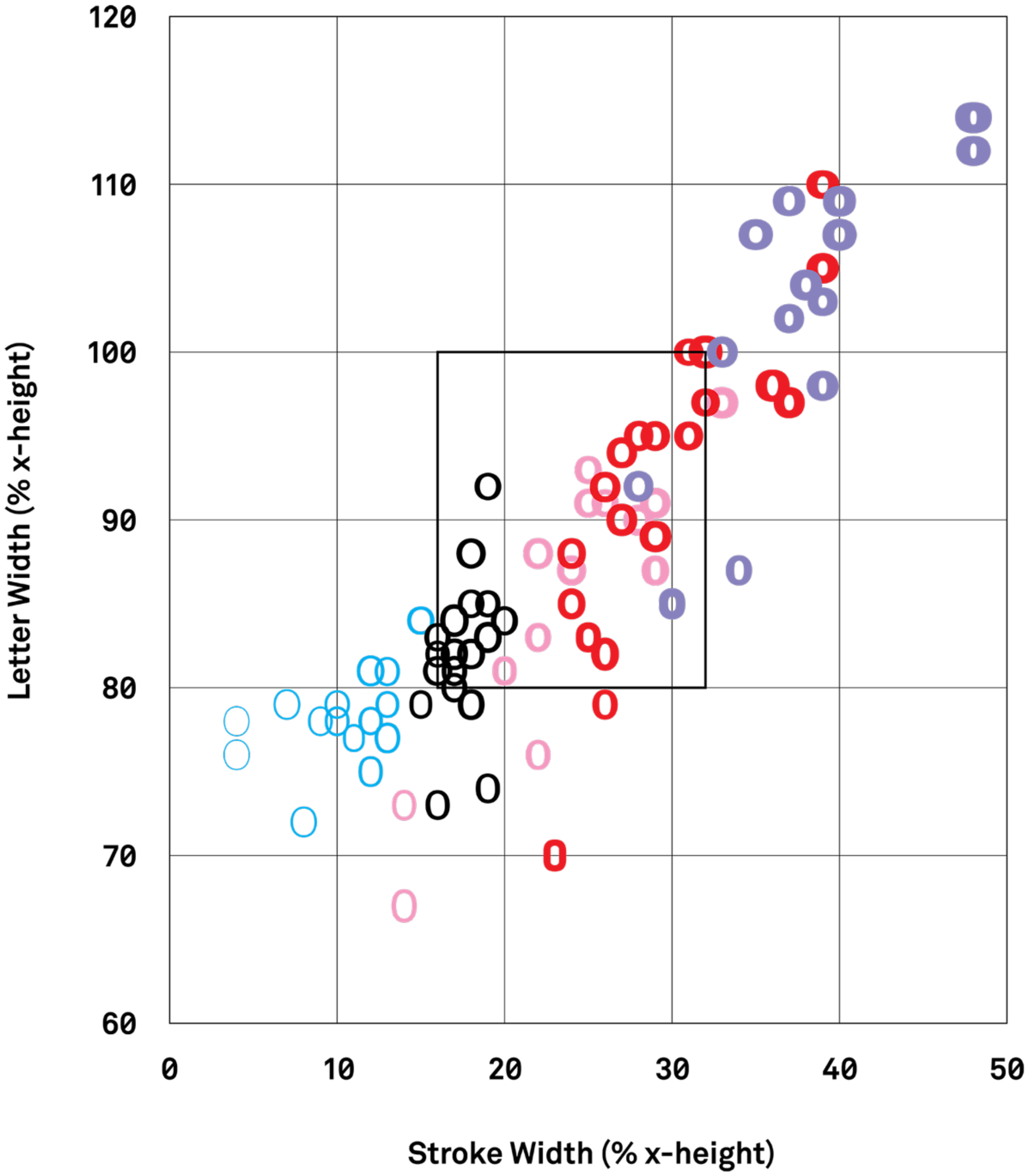


Figure 57: Relationship between the laboratory typeface parameters (black rectangle) and roman text typefaces colour-coded according to weight nomenclature: light (blue), regular (black), medium (pink), bold (red), and extra-bold (violet) typefaces.

The first master is designed to represent a roman regular weight typeface. The master’s stroke width is 16% and its letter width is 80%. The master’s parameters are represented by the bottom left corner of the black rectangle (see Figure 57). Figure 57 illustrates the cluster of roman regular (black) typefaces near the bottom left corner of the rectangle. The parameters of the master represent the lower stroke width (16%) of this cluster of regular weight typefaces, and the corresponding cluster of lower letter widths (80%).

It is useful to note at this point that in choosing the final parameter values of the laboratory typeface, simple ones are selected whenever possible. This is to minimise calculation errors when working in FontLab, where the parameters have to be translated into typeface design units known as UPM (Units Per eM) (FontLab, 2006). For example, the FontLab default x-height is 500 UPM (Cabarga, 2006), therefore a letter width of 80% x-height translates to a tractable 400 UPM letter width.

The relationship between the laboratory typeface parameters (i.e. black rectangle) and roman, extended, and condensed typefaces is visualised in Figure 58. Visualising this relationship is necessary to describe the parameters of the other three masters. The second master’s parameters are represented by the bottom right corner of the black rectangle (see Figure 58). This master’s parameters represent the largest stroke width (32%) found at approximately the same letter width (80%) as the first master. Figure 58 illustrates the black rectangle intersecting a blue coloured ‘o’, which is a condensed bold typeface that informs the parameters of this master.

The third master’s parameters are represented by the top right corner of the black rectangle (see Figure 58). This represents the upper letter width (100%) found at the same stroke width (32%) as the previous master. Figure 57 illustrates that this master represents an upper stroke width of bold typefaces, as the majority of roman bold (red) typefaces have lower stroke width values. As stated, Figure 57 illustrates that the laboratory typeface parameters delineate an area that includes most of the roman regular (black) and bold (red) typefaces.

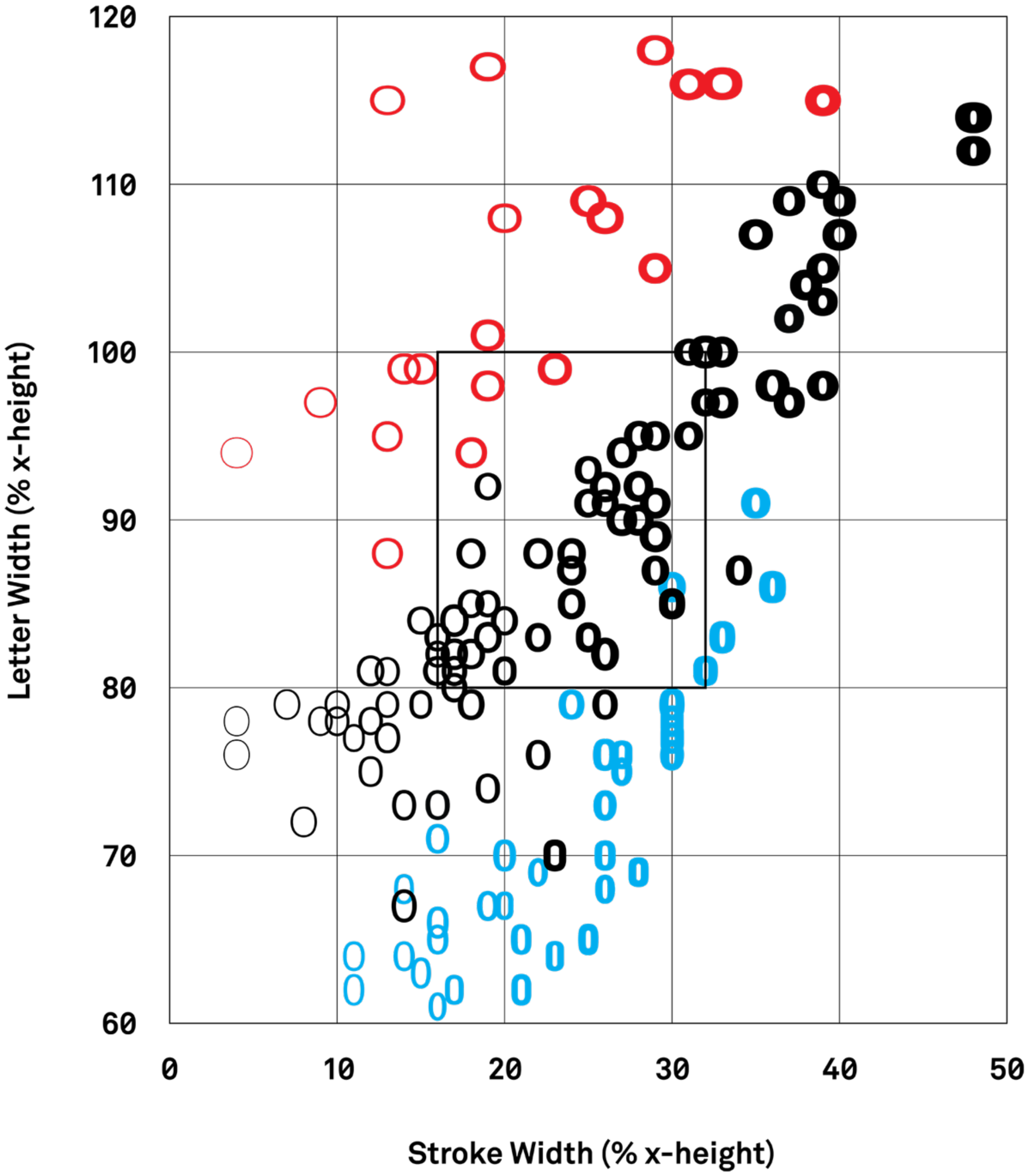


Figure 58: Relationship between the laboratory typeface parameters (black rectangle) and text typefaces colour-coded according to width nomenclature: regular (black), extended (red), and condensed (blue) typefaces.

The fourth and final master's parameters are represented by the top left corner of the black rectangle (see Figure 58). This master's parameters are based on the letter width of the third master (100%) and the stroke width of the first master (16%). Figure 58 illustrates that this master has similar proportions to some of the extended (red) typefaces.

These four masters illustrate that the parameters of the laboratory typeface are not only determined by stroke width values of interest. They fundamentally depend on being able to create a matrix of three stroke widths by three letter widths, in which all nine typefaces reflect the proportions of text typefaces. Figure 59 illustrates the relationship between the laboratory typeface and roman, condensed, and extended typefaces. The laboratory typeface maps well to a large range of roman typefaces, while still reflecting the other extremes by representing extended and condensed typefaces.

The laboratory typeface parameters also reflect scientific knowledge. Figure 60 illustrates the relationship between the laboratory typeface parameters and the sans serif typefaces tested by Shaw (1969), Smither and Braun (1992), and Sheedy et al. (2005). An analysis of these studies reveals three main insights (section 5.3.1) which inform the design of the laboratory typeface. First, the typefaces found to be more legible are all above 22% stroke width. Therefore, the parameters of the laboratory typeface are designed to test stroke widths above and below 22% (Figure 60). Second, the typefaces found to be more legible have stroke widths as high as 31% (Gill Sans Bold) (Shaw, 1969) and 32% (Verdana Bold) (Sheedy et al., 2005). This informs the highest stroke width of the masters, which are designed at 32% (Figure 60). Third, typefaces with stroke widths of 31% (Gill Sans Bold) and 32% (Verdana Bold) have letter widths of 100%. This informs the highest letter width of the masters, which are designed at 100% (Figure 60). Figure 60 illustrates that the highest stroke width (32%) and letter width (100%) of the masters (top right corner of the black rectangle) correspond with the parameters of typefaces found to improve legibility: Gill Sans Bold (31% stroke width, 100% letter width) (Shaw, 1969) and Verdana Bold (32% stroke width, 100% letter width) (Sheedy et al., 2005).

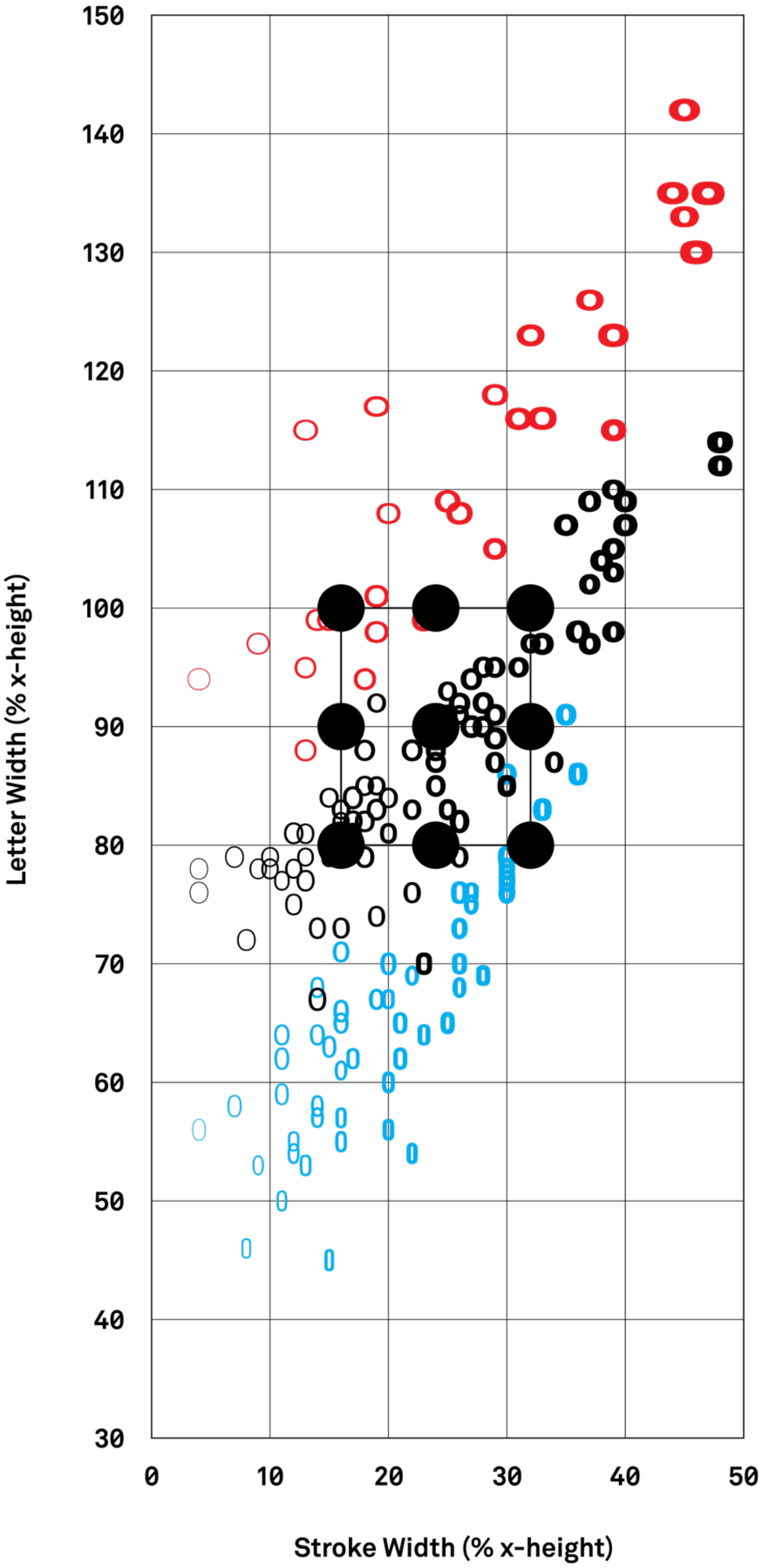


Figure 59: Nine laboratory typefaces represented by black circles, overlapping text typefaces colour-coded according to width nomenclature: regular (black), extended (red), and condensed (blue) typefaces.

Figure 60: Relationship between laboratory typeface parameters (black rectangle) and a subset of scientific studies. List of studies: Shaw, 1969; Smither & Braun, 1994; Sheedy et al., 2005.

The visualisation of design knowledge also informs the analysis of scientific knowledge. As discussed in section 5.3.1, the visualisation of text typeface families (Figure 54) indicates that the 39% stroke width of Franklin Gothic Heavy (Subbaram, 2004) is too high for the laboratory typeface. Specifically, Figure 54 illustrates that most bold typefaces are below 39% stroke width.

5.4 LABORATORY TYPEFACE DESIGN

The development of the laboratory typeface required designing four masters at the extremes of weight and width, and then interpolating five typefaces. Lowercase characters of the four masters are shown in Figure 61. The laboratory typeface family consists of nine individual typefaces (Figure 62), representing a matrix of three stroke widths by three letter widths (Figure 59).

The laboratory typeface is a sans serif for two major reasons. First, as stated in section 5.2.1, evidence indicates that sans serif typefaces improve reading performance for people with low vision (e.g. Morris et al., 2002). Second, sans serif typefaces have lower stroke contrast which is more suitable for a laboratory typeface designed to investigate stroke width.

The main design strategy is to reflect typeface design conventions while controlling parameters as much as possible. The condensed bold master is the most difficult to draw, and in many cases sets the parameters for the other three masters. A strategy of using the fewest number of points to draw each letter is employed, facilitating the design work because points must be in the same relative position in each master to make them compatible. The laboratory typeface is also designed with symmetrical counters. This is appropriate for the condensed bold master, and in order to maintain consistency this is extended through the masters. Figure 63 compares the ‘n’s of the four masters with Helvetica Neue. Note that Helvetica Black Condensed has a symmetrical counter, however the other typefaces in the family do not. This is in contrast to the laboratory typeface masters which each have a symmetrical counter.

Figure 61: Master typefaces: (A) Roman Bold, wt1wd1; (B) Extended Regular, wt0wd1; (C) Condensed Bold, wt1wd0; (D) Roman Regular, wt0wd0.

Figure 62: Nine laboratory typefaces (top to bottom): (A) Letter width 100%, Stroke width 32%, 24%, 16%; (B) Letter width 90%, Stroke width: 32%, 24%, 16%; (C) Letter width 80%, Stroke width: 32%, 24%, 16%.

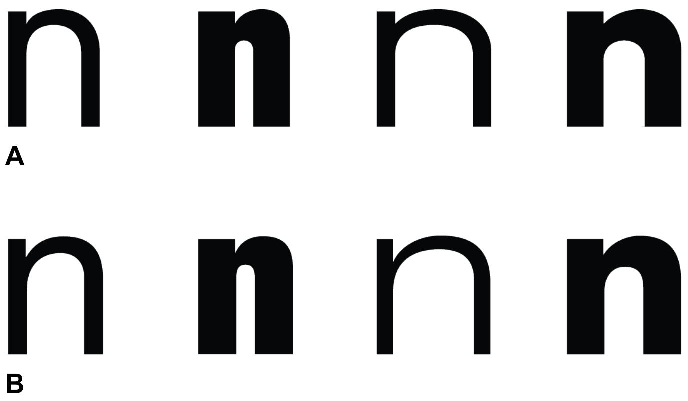


Figure 63: Comparison of (A) The laboratory typeface and (B) Helvetica Neue. Laboratory typeface from left to right: wt0wd0, wt1wd0, wt0wd1, and wt1wd1. Helvetica Neue from left to right: Roman, Condensed Black, Extended Light, and Heavy.

The laboratory typeface is also drawn with squared versus rounded forms. This helps to minimise alignment zones and therefore better control the height of each character. *Alignment zones* set the degree to which letters overshoot the baseline and x-height (Cabarga, 2006), for example the letter ‘o’. In the same way, squared forms minimise the difference in letter width between straight and round edged characters, for example the ‘n’ and ‘o’ respectively.

The letter width is challenging to control, as the relationships between the widths of letters change across typefaces within a family. The width of each letter is calculated as a percentage of the width of the letter ‘n’, and this is kept consistent across the masters. For example, the ‘m’ is 164% of the width of the ‘n’ in each of the masters. Widths are also held constant within groups of letters with similar widths, for example ‘r’ ‘t’ ‘f’ are designed at 70% of the width of the ‘n’, across the masters.

Stroke width is the most important parameter to control. In order to accomplish this, the masters are designed with low contrast. Using the ‘n’ as an example, stroke contrast is calculated as the thickness of the horizontal stroke as a percentage of the thickness of the vertical stroke. Stroke contrast is minimised within masters, held constant across letter widths, and differs slightly across stroke widths. Stroke contrast is 85% for lighter weight masters (i.e. wt0wd0, wt0wd1), and 75% for heavier weight masters (i.e. wt1wd0, wt1wd1). When designing the ‘n’, it is also common for the stroke to narrow at the join, which is minimised in the masters compared to commercial typefaces like Helvetica Neue (see Figure 63). It is also common for stem lengths to differ across typefaces in a family, which is held constant within the masters at 11% of x-height. Figure 63 illustrates the varying stem lengths in the Helvetica Neue typeface family.

5.5 DISCUSSION

5.5.1 LABORATORY TYPEFACE AND CHAPTER 4 SCIENTIFIC REVIEW

While the laboratory typeface was designed in 2009, it is aligned with the full review of scientific studies presented in Chapter 4. Figure 64 illustrates the relationship between the laboratory typeface parameters and the six scientific studies reviewed in Chapter 4 which form the basis of the recommendations (section 4.4.2) (Roethlein, 1912; Shaw, 1969; Smither & Braun, 1994; Sheedy et al., 2005; Beveratou, 2016; Beier & Oderkerk, 2019a). The review in Chapter 4 recommends examining stroke widths of both intermediate (i.e. between 22-29%) and higher (i.e. above 30%) values to determine an optimum for low vision readers. The laboratory typeface’s parameters align with these recommendations. This is because two of the studies informing the recommendations in 2021 (Shaw, 1969; Sheedy et al., 2005) were also central to decision making in 2009.

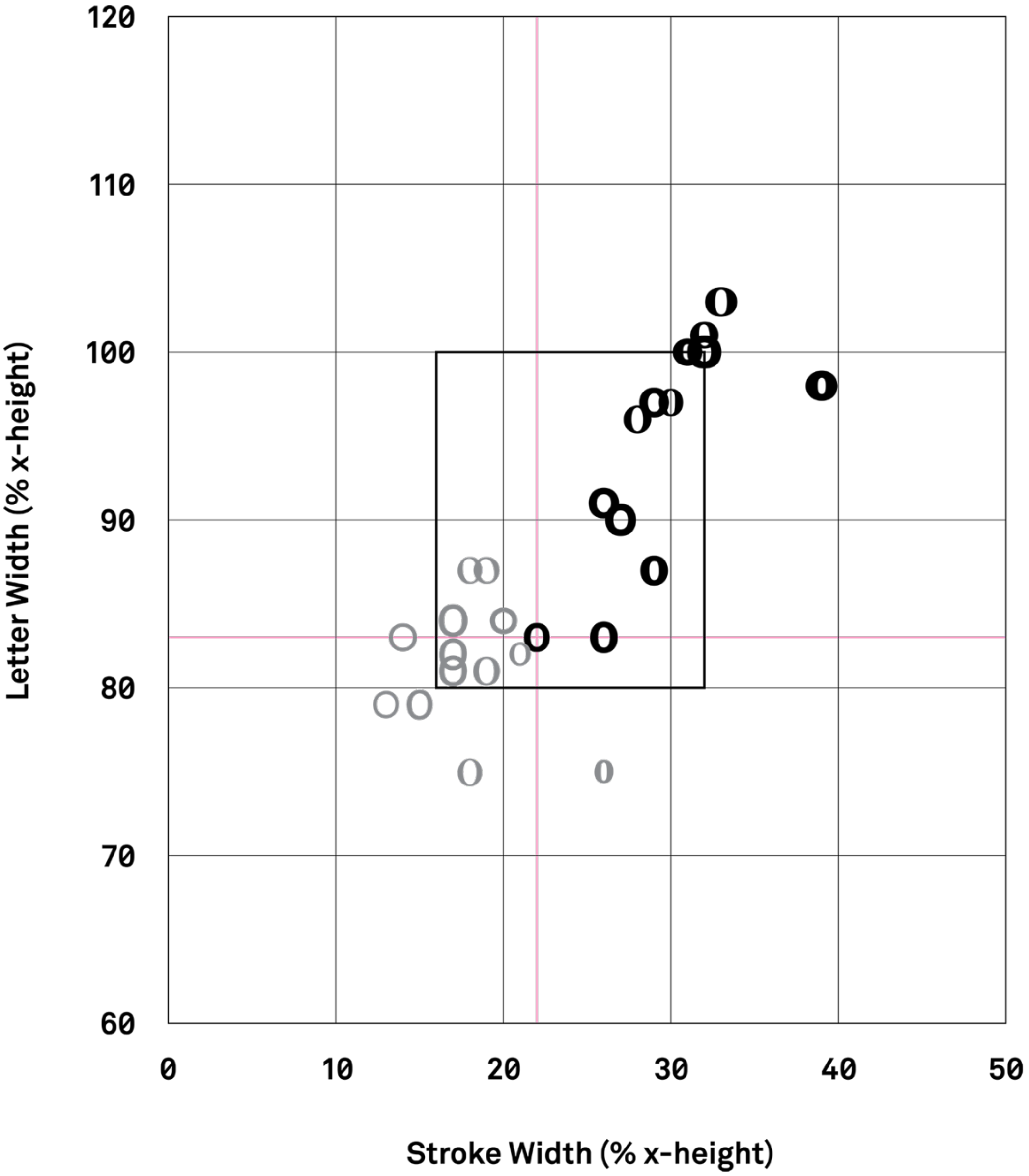


Figure 64: Relationship between laboratory typeface parameters (black rectangle) and six scientific studies reviewed in Chapter 4. List of studies: Roethlein, 1912; Shaw, 1969; Smither & Braun, 1994; Sheedy et al., 2005; Beveratou, 2016; Beier & Oderkerk, 2019a. The upper right quadrant contains typefaces found to have higher legibility.

5.5.2 LABORATORY TYPEFACE AND PHD DIRECTION

Instead of being used as test material to address the research question experimentally, the laboratory typeface serves a different purpose within the PhD research. First, the typeface offers an exemplar, illustrating that experimental test material can be created that allows for the controlled study of stroke width, while still reflecting design practice. As described in section 5.4, the typeface design process requires many decisions and compromises, in both controlling parameters and reflecting conventional letterform construction. A design-informed laboratory typeface for the controlled study of stroke width is yet to be published in the research literature.

Second, and more importantly, the typeface design process gives rise to the quantitative analysis of typefaces; the measurement and visualisation of typeface proportions. This influences the change in PhD direction toward the consolidation of scientific knowledge, generation of design knowledge, and development of interdisciplinary knowledge through information visualisation. In response to this change in research direction, the consolidation of scientific knowledge becomes larger in scope, encompassing serif typefaces as presented in Chapter 4. Further, while the design knowledge generated within Chapter 5 (section 5.2.2) is sufficient to create the laboratory typeface, Chapter 6 presents a more objective investigation. A points-based survey of design sources (e.g. typeface best-sellers lists) is employed to determine a set of typefaces to serve as the basis for the generation of design knowledge (section 6.2). This design knowledge is a contribution unto itself, however its more important role is in the development of interdisciplinary—science and design—typographic knowledge.

5.6 SUMMARY

This chapter describes the creation of a laboratory typeface based on scientific and design knowledge. Scientific knowledge informs the choice of parameters, such that the typeface is capable of testing the stroke widths of typefaces found to improve reading performance. Design knowledge informs the choice of parameters, such that the typeface reflects the proportions of text typeface families. Instead of being used for experimental study, the laboratory typeface offers an exemplar illustrating that test material can be developed for the controlled study of stroke width while reflecting design practice. More importantly, the typeface design process gives rise to the quantitative analysis of typefaces, playing a crucial role in the final direction of the PhD research.

Chapter 5 illustrates how the quantitative analysis of typefaces—measurement and visualisation—is employed to develop a laboratory typeface. Next, Chapter 6 shows how this approach is used to generate interdisciplinary knowledge.

CHAPTER 6. GENERATING DESIGN AND INTERDISCIPLINARY KNOWLEDGE

6.1 INTRODUCTION

This chapter is focused on generating design knowledge through a design phenomenology study, focused on sans serif text typeface proportions. By integrating this knowledge with the consolidation of scientific knowledge in Chapter 4, the following question can be addressed: How do the proportions of typefaces found to have higher legibility relate to typefaces commonly used in design practice? The main outcome of this chapter is the development of this interdisciplinary knowledge on typeface legibility for low vision readers. This addresses the research question in the context of design practice, investigating optimal typeface weights for low vision adults.

This chapter is presented in two parts. Part 1 describes a points-based survey of design sources, undertaken to determine twenty sans serif text typefaces representative of design practice. This objective method is in contrast to the choice of typefaces based on professional knowledge in Chapter 5 (section 5.2.2). The survey was initially conducted in 2010 and then updated in 2020. My intention is that the final typefaces chosen as the basis for the generation of design knowledge have an established record within the design community.

In Part 2 of this chapter, twenty typefaces are measured and visualised, and summary statistics are reported to support the analysis. The visualisations analyse and communicate the stroke width and letter width of sans serif text typeface families. Weight nomenclature is visualised, illustrating the proportions of regular and bold typefaces. Statistics including the median stroke width values for regular and bold typefaces are reported. Finally, visualisations of interdisciplinary knowledge illustrate the relationship between the proportions of typefaces tested experimentally and sans serif text typefaces. The implications for inclusive print guidelines are addressed in the discussion.

6.2 PART 1: SELECTION OF TYPEFACES TO INVESTIGATE

6.2.1 METHODS

6.2.1.1 SURVEY OF DESIGN SOURCES AND INCLUSION CRITERIA: 2010

The 2010 investigation examines multiple design sources in order to determine which typefaces are relevant to typography practice. In total, twenty-three sources contribute to the dataset: five typeface distributors, nine typography manuals, seven history books, and two ‘top 100’ lists. Typefaces are allotted a maximum of one point per source that they appear in, resulting in total points per typeface ranging from one to twenty-three. These points are used to determine a group of sans serif typefaces which are analysed to generate design knowledge. Note that the purpose of this survey is not to rank typefaces, but to arrive at a group of typefaces that are relevant to the practicing design community. Appendix 4 lists the final twenty typefaces alongside the number of points that are accrued from the four types of sources.

Each type of design source captures a different aspect of a typeface’s importance. Typeface distribution sales are a direct assay of use by designers. Typography manuals highlight typefaces that experienced designers deem as central to the practice, and again reference contemporary use. History books on the other hand, represent typefaces of historical importance. This source is still relevant to contemporary practice, as many of the most popular typefaces have a relatively long history of use (e.g. Helvetica). Two “top 100” typeface lists are also utilised in the investigation; one using sales, historical importance, and aesthetics as its criteria (FontShop, 2022), and the other emphasising historical, technological, and theoretical importance (Shaw, 2010).

The best-selling typefaces of five internationally renowned type distributors are investigated: fifteen from Adobe Fonts (2022); fifty from FontShop

(2020); twenty from Linotype (2020); twenty-five from Monotype (Fonts.com, 2020a), and fifty from MyFonts (2020). The number of typefaces that each distributor contributes to the dataset is based on the number of typefaces presented on their websites, with the exception of FontShop (2020) which displays two hundred best-selling typefaces and is capped at fifty. A typeface family can only receive a maximum of one point per distributor, therefore if two typefaces from the same family are both best sellers, only one point is allotted. Distributers excluded from the dataset include smaller type foundries (e.g. Lineto), those known for display type (e.g. T26), those selling through a major distributor (e.g. FontFont sold by FontShop), and those without a best-sellers list (e.g. Berthold).

The typography manuals investigated date from 2002 to 2010. Within each book, the section that focuses on typeface classification is examined. In order to receive a point, a typeface must be presented as a figure. Therefore, if a typeface is mentioned in the text without a visual representation, it is not allotted a point. While this strategy is efficient, it is also based on the assumption that the most important typefaces are presented visually. In total, nine typography manuals are investigated: Ambrose and Harris (2005); Baines and Haslam (2005); Bringhurst (2002); Carter, Day and Meggs (2007); Craig, Scala and Bevington (2006); Ellison (2006); Hill (2010); Jury (2004); Lupton (2004).

Six graphic design history books are investigated: Cramsie (2010); Eskilson (2007); Hollis (2001); Meggs (1998); Poynor (2003); Purvis and Le Coultre (2003). One typeface history book is also included: Dodd (2006). The history books are examined in their entirety, and a typeface is allotted a point if it is presented in a figure.

Two ‘top 100’ typeface lists are investigated. A typeface is allotted one point for each list it appears on, accumulating a maximum of two points. The *Top 100 Typefaces of all Time* (Devroye, 2022) was originally compiled by Paul Shaw in 1998 for Letterspace, the newsletter of the Type Directors Club (Shaw, 2010). Shaw is a type historian and based his list on the historical, technological, or theoretical importance of each design rather than its popularity or aesthetics (Shaw, 2010).

The second list was compiled by FontShop in 2007: *Die 100 Besten Schriften Aller Zeiten (The 100 Best Typefaces of All Time)* (FontShop, 2022). Their jury consisted of Roger Black, Stephen Coles, Jan Middendorp, Veronika Elsner, Ralf Herrmann, Bertram Schmidt-Friderichs, and Claudia Guminski. The judging was based on three criteria: sales (40%), historical importance (30%), and aesthetic qualities (30%). The first criterion represents objective data based on contemporary usage, while the latter two criteria are more subjective.

The 2010 investigation results in a list of typefaces, each with an allotted number of points according to the sources described. Typefaces are then removed from the dataset in order to focus on typefaces designed for the contemporary and conventional practice of setting continuous text. Typefaces that are excluded from further analysis are those classified as script, italic, blackletter, monospaced, postmodern, display, or decorative. Typefaces with only uppercase characters, not designed for human reading, or lacking a digital equivalent are also removed. Further, styles without specific reference to a particular typeface (e.g. Egyptian) are removed. Typefaces created from 2002 onward are also removed from the dataset in order to ensure that all typefaces under investigation have a longstanding history of use in the practicing design community and are not representing a potentially short-lived period of popularity. Lastly, serif typefaces are removed because this investigation is focused on sans serif typefaces.

Points primarily determine the final inclusion of typefaces within the 2010 dataset, with one exception. As the dataset is intended to represent contemporary design practice, typefaces with over 50% of their points accrued through history books are excluded. Final inclusion in the 2010 dataset occurs if a typeface meets the above criteria and has accrued four or more points, resulting in twenty-seven typefaces in total.

6.2.1.2 SURVEY OF DESIGN SOURCES AND INCLUSION CRITERIA: 2020

The 2010 dataset is integrated with a 2020 investigation in order to ensure that the final list of typefaces aligns with design practice in 2020. In order to accomplish this, data is gathered from the typeface distributers examined in section 6.2.1.1 that continue to publish best-sellers lists. The top one hundred best-selling typefaces from FontShop (2020), Linotype (2020), and Monotype (Fonts.com, 2020a), and the top fifty from MyFonts (2020) are recorded. Note that Monotype (Fonts.com, 2020a) and MyFonts (2020) publish lists of 100 and 50 best-selling typefaces, respectively, while FontShop (2020) and Linotype (2020) publish lists that extend via scrolling and are capped at 100 best-selling typefaces.

Inclusion in the final dataset requires that a typeface meets the inclusion criteria for the 2010 dataset as well as appearing on any one of the four distributors’ best-sellers lists in 2020. While only one point accrued from 2020 best-sellers lists is required, total points recorded per typeface family in 2020 are reported in Appendix 4. By combining the 2010 investigation with the 2020 investigation, my intention is that the final list of typefaces have an established record of use.

A total of twenty-four typefaces meet the above criteria. Typefaces are then removed from the dataset based on final criteria regarding their recommended usage. Four typefaces classified as geometric sans are removed; Kabel based on being recommended for title and logo work (Fonts.com, 2020e), and Avant Garde Gothic (Fonts.com, 2020d), Eurostile (Fonts.com, 2020b), and Futura (Fonts.com, 2020c) for being recommended for setting small amounts of text. The final dataset is comprised of twenty sans serif typeface families recommended for the setting of continuous text.

6.2.2 RESULTS: TWENTY SANS SERIF TEXT TYPEFACE FAMILIES

The final twenty typefaces included in the dataset are presented in Table 16. Figure 65 shows the lowercase characters of the twenty sans serif typefaces investigated. Figure 66 shows the bold weight of the same twenty typefaces.

Table 16: Twenty sans serif typefaces investigated. The ‘x’ indicates whether roman, condensed, and extended typefaces are analysed.

Typeface	Roman	Condensed	Extended
Arial	x	x	
Avenir	x		
DIN	x	x	
Franklin Gothic	x	x	
Frutiger	x	x	
Gill Sans	x	x	
Helvetica	x	x	
Helvetica Neue	x	x	x
Meta	x	x	
Myriad	x	x	
News Gothic	x	x	
Officina Sans	x		
Optima	x		
Rotis Sans	x		
Scala Sans	x	x	
Stone Sans	x		
Syntax	x		
TheSans	x		
Trade Gothic	x	x	x
Univers	x	x	x

Arial	abcdefghijklmnopqrstuvwxyz
DIN	abcdefghijklmnopqrstuvwxyz
Avenir	abcdefghijklmnopqrstuvwxyz
Franklin Gothic	abcdefghijklmnopqrstuvwxyz
Frutiger	abcdefghijklmnopqrstuvwxyz
Gill Sans	abcdefghijklmnopqrstuvwxyz
Helvetica	abcdefghijklmnopqrstuvwxyz
Helvetica Neue	abcdefghijklmnopqrstuvwxyz
Meta	abcdefghijklmnopqrstuvwxyz
Myriad	abcdefghijklmnopqrstuvwxyz
News Gothic	abcdefghijklmnopqrstuvwxyz
Officina Sans	abcdefghijklmnopqrstuvwxyz
Optima	abcdefghijklmnopqrstuvwxyz
Rotis Sans	abcdefghijklmnopqrstuvwxyz
Scala Sans	abcdefghijklmnopqrstuvwxyz
Stone Sans	abcdefghijklmnopqrstuvwxyz
Syntax	abcdefghijklmnopqrstuvwxyz
The Sans	abcdefghijklmnopqrstuvwxyz
Trade Gothic	abcdefghijklmnopqrstuvwxyz
Univers	abcdefghijklmnopqrstuvwxyz

Figure 65: Twenty roman regular sans serif typefaces investigated. See section 6.3.1.2 for further details regarding weight names.

Arial	abcdefghijklmnopqrstuvwxyz
DIN	abcdefghijklmnopqrstuvwxyz
Avenir	abcdefghijklmnopqrstuvwxyz
Franklin Gothic	abcdefghijklmnopqrstuvwxyz
Frutiger	abcdefghijklmnopqrstuvwxyz
Gill Sans	abcdefghijklmnopqrstuvwxyz
Helvetica	abcdefghijklmnopqrstuvwxyz
Helvetica Neue	abcdefghijklmnopqrstuvwxyz
Meta	abcdefghijklmnopqrstuvwxyz
Myriad	abcdefghijklmnopqrstuvwxyz
News Gothic	abcdefghijklmnopqrstuvwxyz
Officina Sans	abcdefghijklmnopqrstuvwxyz
Optima	abcdefghijklmnopqrstuvwxyz
Rotis Sans	abcdefghijklmnopqrstuvwxyz
Scala Sans	abcdefghijklmnopqrstuvwxyz
Stone Sans	abcdefghijklmnopqrstuvwxyz
Syntax	abcdefghijklmnopqrstuvwxyz
The Sans	abcdefghijklmnopqrstuvwxyz
Trade Gothic	abcdefghijklmnopqrstuvwxyz
Univers	abcdefghijklmnopqrstuvwxyz

Figure 66: Twenty roman bold sans serif typefaces investigated. See section 6.3.1.2 for further details regarding weight names.

6.3 PART 2: GENERATING DESIGN AND INTERDISCIPLINARY KNOWLEDGE

6.3.1 METHODS

6.3.1.1 VISUALISING TYPEFACE DATA

The measurement of typeface proportions employs similar methods as presented in Chapter 4 (section 4.2.2). One exception is that typeface data is recorded to three significant digits (e.g. 13.5% stroke width). This serves to decrease the chances that two typefaces have identical data. In Chapter 5, whenever two typefaces have identical stroke width and letter width values—resulting in two ‘o’s overlapping when visualised—one ‘o’ is deleted. This occurs in twelve instances (section 5.2.3). While the visualisations of design knowledge in Chapter 5 communicate the proportions of typefaces utilised in design practice (e.g. Figure 57), the number of typefaces with those proportions is not communicated. This is sufficient to inform the design of a laboratory typeface that reflects design knowledge.

In contrast, Chapter 6 is focused on generating design knowledge based on twenty sans serif typefaces determined to be representative of design practice. Visualising the proportions of each of the twenty typefaces is important for the comprehensiveness of this design phenomenology study. The result of recording data to three significant digits is observable in Figure 71b, which shows a cluster of three ‘o’s in the lower left quadrant. If typeface data is recorded to two significant digits, Arial, Helvetica, and Syntax Regular have identical stroke width (17%) and letter width (81%) values. However, their proportions are distinct when recorded to three significant digits, for example stroke width values of 17.0%, 16.8%, and 16.9% for Arial, Helvetica, and Syntax Regular, respectively. As such, each of these typefaces’ data can be visualised simultaneously.

Typeface data is visualised using similar methods as presented in Chapters 4 (section 4.2.3) and 5 (section 5.2.3), with three exceptions. First, the graph is redesigned; standard axes are replaced by rectangles which visually represent stroke width and letter width proportions (e.g. Figure 67). This is done to provide the viewer with further visual references for stroke width and letter width values, with the intention of making the graph easier to understand.

The second difference in visualisation methods is that plotted typeface ‘o’s are scaled to an equivalent x-height, which serves to remove a variable that interferes with a more detailed understanding of the visualisations. As in Chapter 4, the positioning of the data points (i.e. visualising stroke width and letter width values) remains of central importance (section 4.2.3). However in Chapter 6, the visualisations are refined to allow for a clearer communication of letterforms. Specifically, the reader is able to more easily compare the letterforms of typefaces experimentally found to improve legibility and typefaces used in design practice (e.g. Figure 71a and 71b).

The third difference in methods is that visualisations which include the four quadrants (section 4.4.2) show the upper right quadrant shaded in pink. This aids in the visualisation of interdisciplinary—scientific and design—knowledge (Figure 71a and 71b), which shows the relationship between typefaces found to have higher legibility (scientific knowledge) and regular weight typefaces commonly used for setting text (design knowledge). As these two groups of typefaces are represented by the colour black in visualisations presented in Chapters 4 and 5 respectively, the pink upper quadrant differentiates the typefaces found to have higher legibility (scientific knowledge). Shading the upper right quadrant also benefits the discussion of further visualisations (Figures 72 and 73).

6.3.1.2 SUMMARY STATISTICS AND TYPEFACES ANALYSED

The analysis of stroke width and letter width values of sans serif text typefaces includes the reporting of summary statistics. Analyses include measures of central tendency and measures of dispersion. A *measure of* *central tendency*, or *average*, is a description of the general “preponderance of values somewhere around the middle of the range of observed values” (Zar, 1999, p.20). Measures of central tendency include the mean and median. The measure of central tendency employed within the thesis is the median. The *median* is defined as “the middle measurement in an ordered set of data” (Zar, 1999, p.23). In other words, “it is the value such that half the numbers in the data set are larger and half are smaller” (Hand, 2008, p. 29). The median (versus mean) is reported because it is not as affected by extreme values, and is a “resistant” statistic (Zar, 1999, p.25). A *measure of dispersion* is an indication of how widely dispersed the data are around the average (Hand, 2008). The range as well as the minimum and maximum are reported for parameters. All statistics are reported to three significant figures.

Summary statistics are reported for roman regular and bold typefaces. As typeface nomenclature is not standardised, these analyses are not possible to conduct using all twenty typefaces. Summary statistics of roman regular typefaces are calculated based on sixteen typefaces with a regular weight (weight names include Regular, Roman, and Plain). Summary statistics are calculated for roman bold typefaces based on eighteen typefaces with a bold weight (weight names include Bold only). Table 17 lists the typefaces that are analysed to produce the summary statistics (see Figures 65 and 66).

Visualisations including all twenty typefaces are accomplished by choosing an appropriate weight within typeface families that are not included in the statistical analysis. These typefaces are chosen by selecting a weight closest to the median stroke width calculated for roman regular and bold typefaces. Visualisations of roman regular typefaces include the sixteen regular weight typefaces (see Table 17), in addition to Franklin Gothic Book, Meta Normal, Officina Sans Book, and Stone Sans Medium. Note that Stone Sans Medium is the lowest weight offered for this typeface family. Visualisations of roman bold typefaces include the eighteen bold weight typefaces (see Table 17), as well as Avenir Heavy and Franklin Gothic Demi.

Table 17: Typefaces analysed to produce summary statistics for roman regular and bold typefaces. The ‘x’ indicates whether the typeface is included in the dataset analysed.

All Typefaces	**Typefaces Analysed to Produce Summary Statistics**	
	Roman Regular Typefaces	**Roman Bold Typefaces**
Arial	x	x
Avenir	x	
DIN	x	x
Franklin Gothic		
Frutiger	x	x
Gill Sans	x	x
Helvetica	x	x
Helvetica Neue	x	x
Meta		x
Myriad	x	x
News Gothic	x	x
Officina Sans		x
Optima	x	x
Rotis Sans	x	x
Scala Sans	x	x
Stone Sans		x
Syntax	x	x
TheSans	x	x
Trade Gothic	x	x
Univers	x	x

6.3.2 RESULTS

6.3.2.1 DESIGN KNOWLEDGE

Figure 67 visualises the stroke width and letter width values of twenty sans serif typeface families. This visualisation represents width categories including roman, extended, and condensed (Table 16), illustrating once again the vast diversity in typeface proportions (as first shown in section 5.3.2). Figure 68 visualises roman typefaces which represent a distinctly smaller range of parameter values. Finally, Figure 69 visualises roman regular typefaces, representing the much smaller range of parameter values of typefaces that are used to set continuous text.

Figure 70 visualises typeface weight nomenclature, illustrating the proportions of roman regular (black) and bold (red) typefaces. These weights are chosen because inclusive typography guidelines recommend a range of weights from regular to bold (e.g. RNIB & ISTD, 2007). The summary statistics presented in Table 18 support the understanding of this visualisation. The analyses indicate that the median stroke width for regular typefaces is 16.7% and for bold typefaces is 26.1% (Table 18). Regular typefaces range from 13.5-19.8% stroke width and bold typefaces from 18.9-40.0% (Table 18). The analyses indicate that while regular and bold typefaces both represent a range of parameter values, bold typefaces represent a larger range in both stroke width and letter width (Table 18). Figure 70 illustrates that while the stroke widths of regular and bold typefaces are generally distinct, Rotis Sans Bold crosses into the range of stroke width values of regular typefaces due to its notably low stroke width (18.9%). As addressed in section 5.3.2, the range of stroke width values of regular (black) and bold (red) typefaces illustrates the lack of standardisation in typeface weight nomenclature. Implications for inclusive print guidelines are considered in the discussion (section 6.4.2).

Table 18: Summary statistics of roman regular and bold sans serif typefaces. All parameter values are given as a percentage of x-height.

Weight	***n***	**Parameter**	**Minimum**	**Maximum**	**Range**	**Median**
Regular	16	Stroke Width	13.5%	19.8%	6.3%	16.7%
	16	Letter Width	71.6%	88.4%	16.8%	82.2%
Bold	18	Stroke Width	18.9%	40.0%	21.1%	26.1%
	18	Letter Width	77.9%	108.8%	30.8%	89.0%

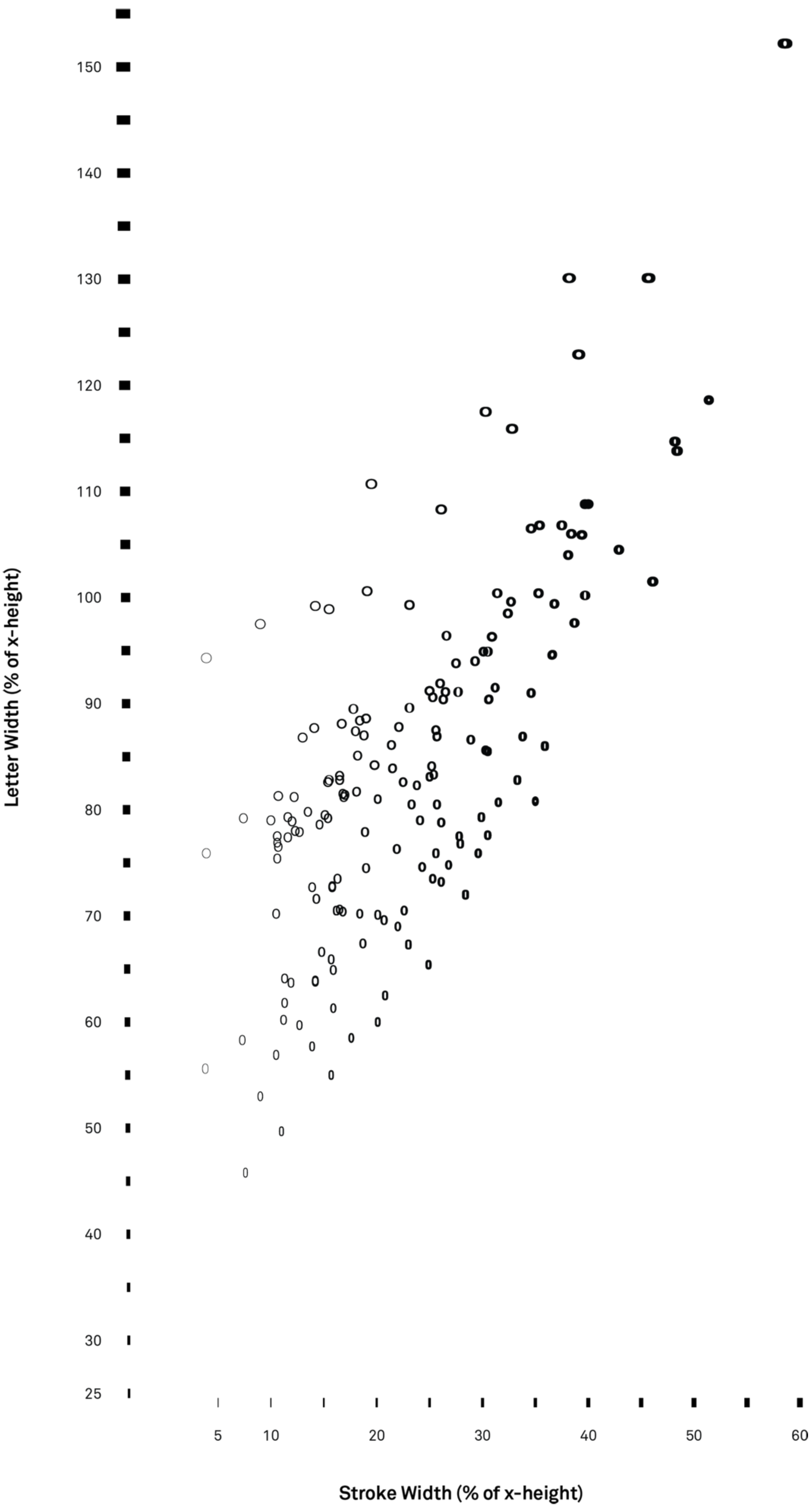


Figure 67: Sans serif roman, extended, and condensed typefaces plotted according to stroke width and letter width.

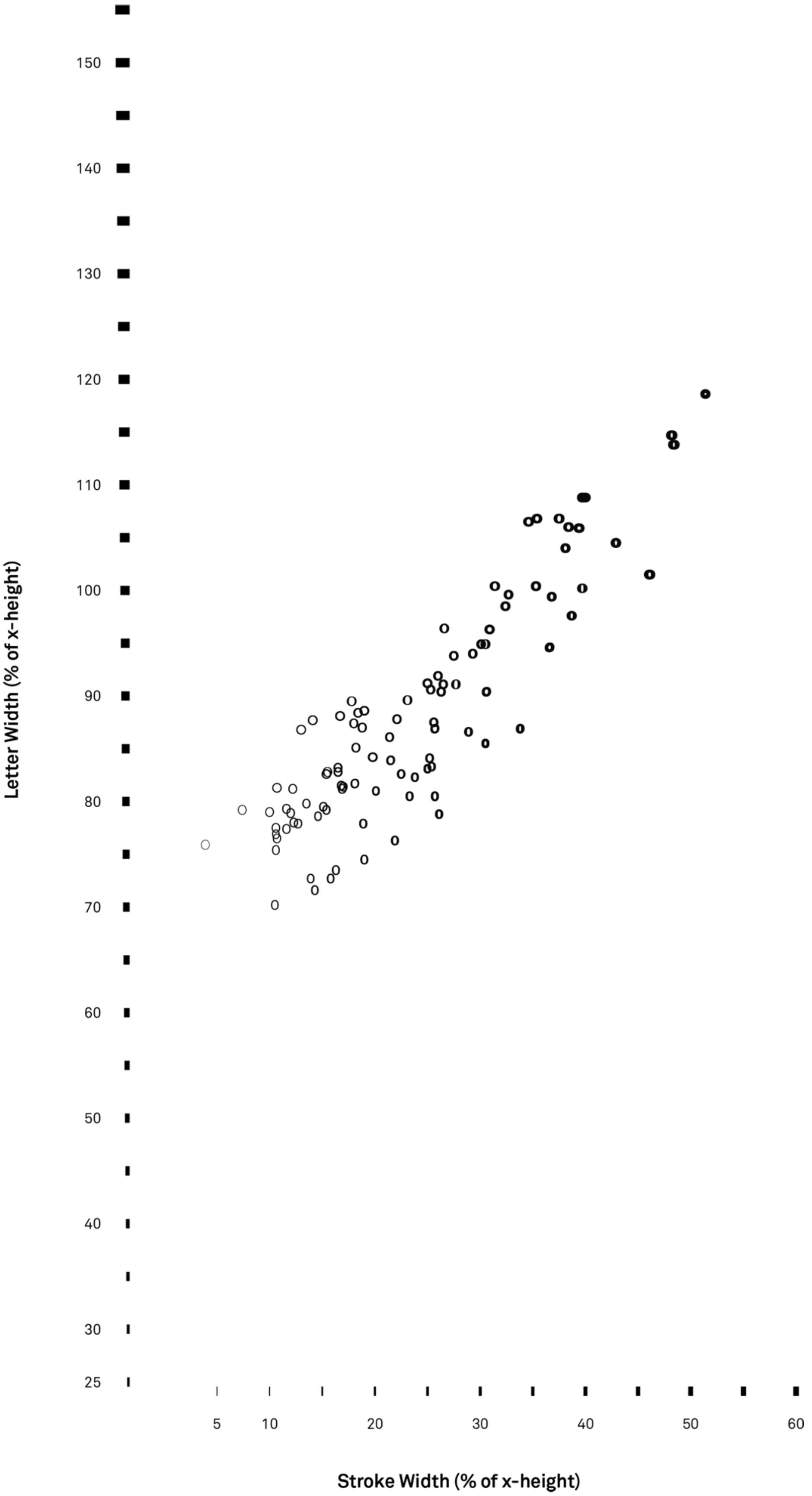


Figure 68: Sans serif roman typefaces plotted according to stroke width and letter width.

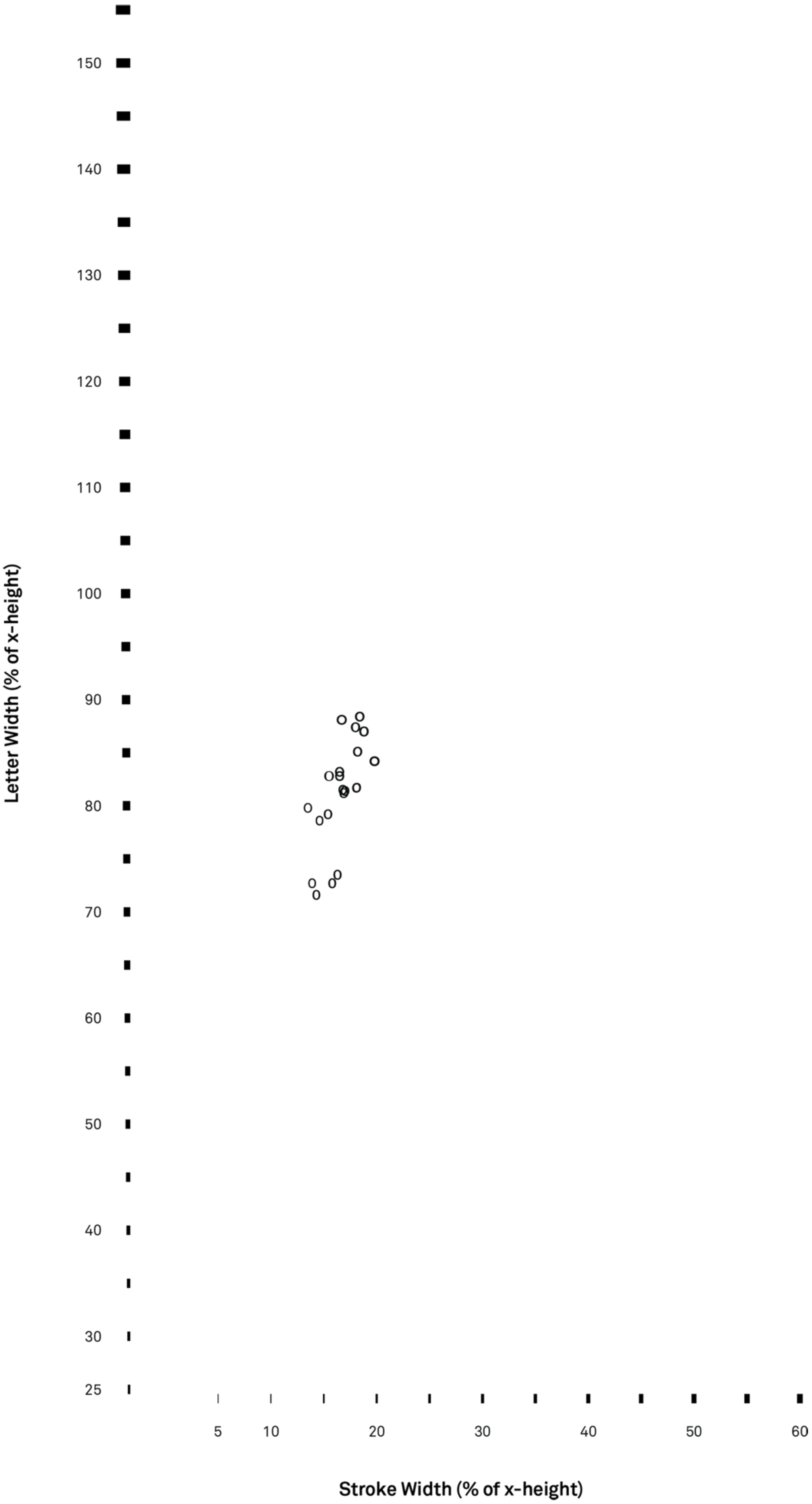


Figure 69: Sans serif roman regular weight typefaces plotted according to stroke width and letter width.

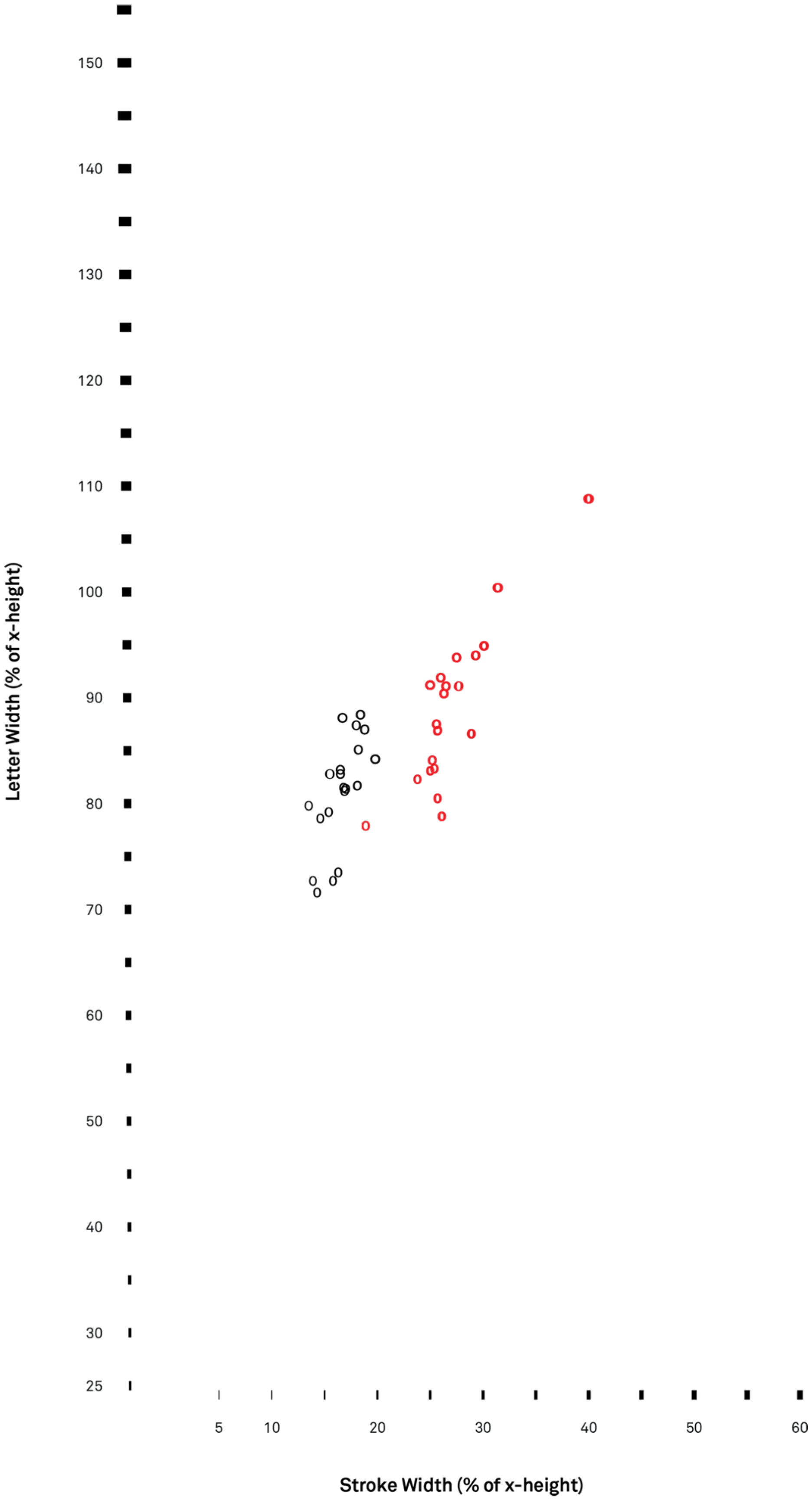


Figure 70: Sans serif weight nomenclature. Typefaces colour-coded to represent regular (black) and bold (red) typefaces.

6.3.2.2 INTERDISCIPLINARY KNOWLEDGE

Figure 71a is a representation of interdisciplinary knowledge, visualising the proportions of typefaces found to have higher legibility (scientific knowledge), alongside twenty roman regular text typefaces (design knowledge). The consolidation of scientific knowledge (section 4.4.2) is based on six scientific studies: Roethlein (1912), Shaw (1969), Smither and Braun (1994), Sheedy et al. (2005), Beveratou (2016), and Beier and Oderkerk (2019a). Figure 71a visualises the same quadrants as Figure 49 in Chapter 4, with the typefaces found to have higher legibility within the upper right quadrant (pink area). Figure 71b is a detail of Figure 71a.

The analysis undertaken in Chapter 4 (section 4.4.2) indicates that a stroke width of 22% improves legibility for normally sighted people reading at threshold type sizes, based on the study by Sheedy et al. (2005). The analysis further indicates that stroke widths higher than 30% improve legibility for adults with low vision and normally sighted people reading at threshold type sizes, based on the studies by Shaw (1969) and Sheedy et al. (2005), respectively. Figure 71a illustrates that all roman regular typefaces have stroke widths below 22%, and none of them exist within the upper right quadrant.

Regarding letter width, Figure 71a illustrates that all typefaces in the upper right quadrant not only have stroke width values of 22% and above, they also have letter width values of 83% and above. This letter width minimum is aligned with the proportions of Franklin Gothic Medium (visualised at the intersection of the two pink lines) and Tiresias LPfont (stroke width 26%, letter width 83%). Figure 71a illustrates that roman regular typefaces not only have stroke widths below 22%, many also have letter widths below 83%, and some dramatically so. The cluster of four ‘o’s below 75% letter width represent Officina Sans, Meta, Trade Gothic, and Rotis Sans.

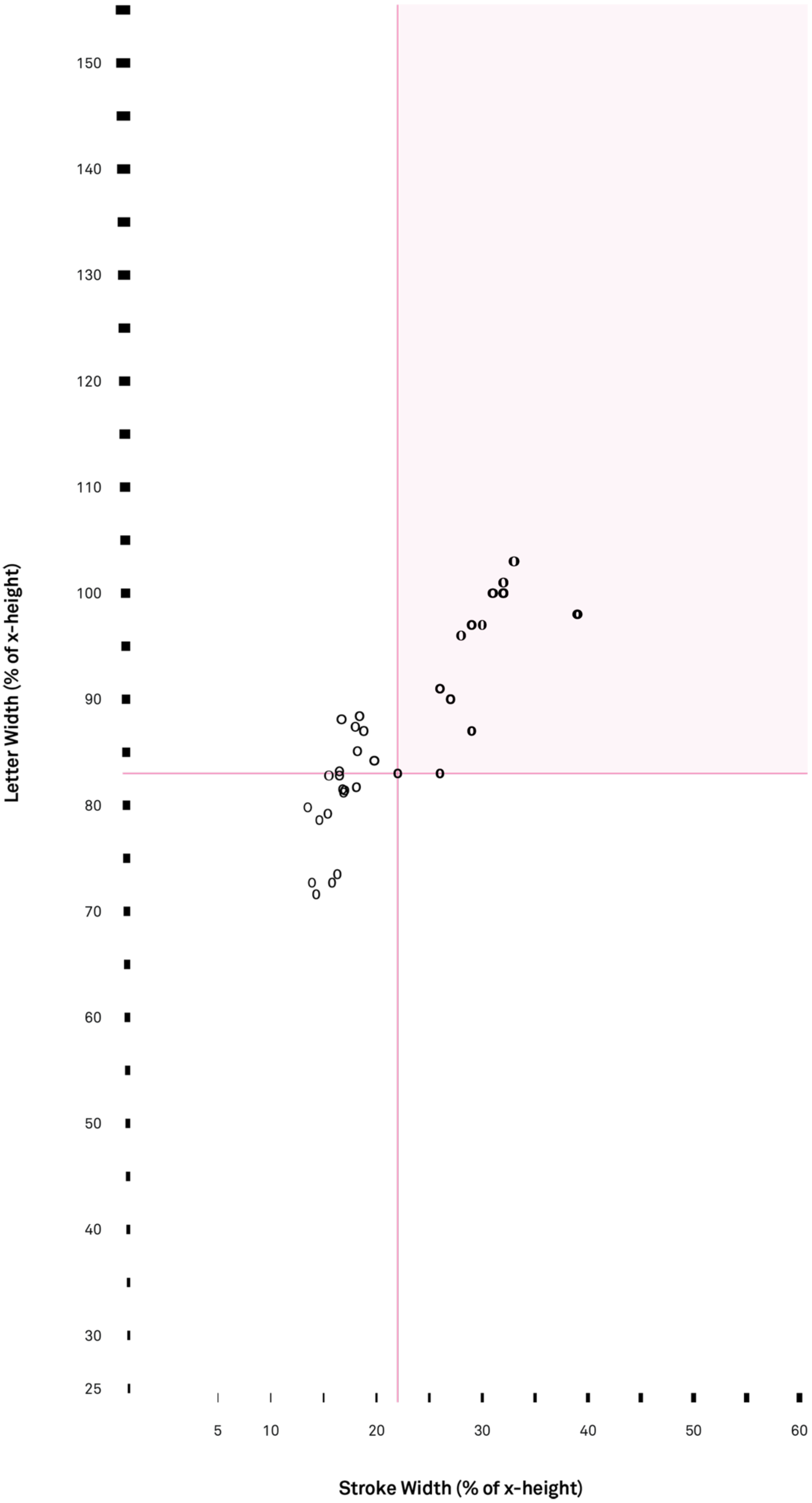


Figure 71a: Relationship between typefaces experimentally found to be more legible (upper right quadrant, pink area) and regular typefaces.

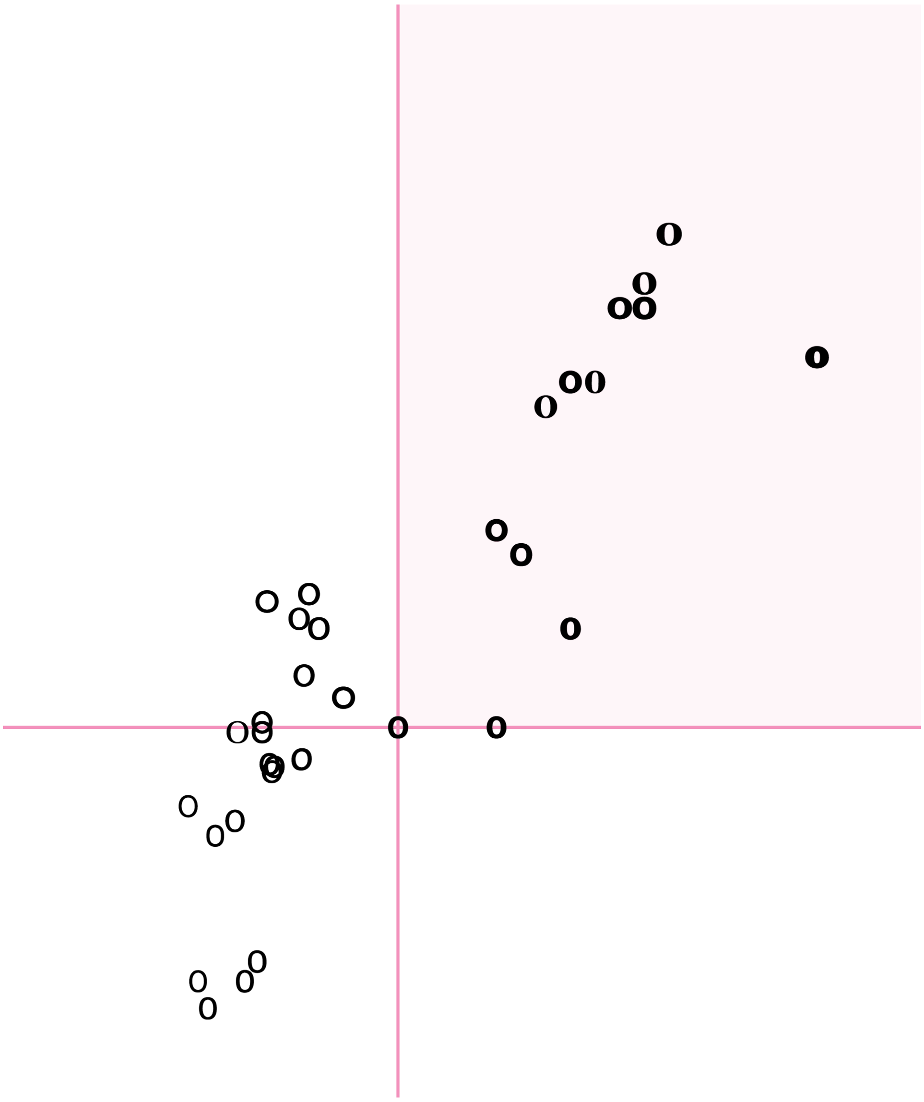


Figure 71b: Detail of figure 71a.

Figure 72 visualises the proportions of typefaces experimentally found to have higher legibility with roman bold typefaces. In comparison to the roman regular typefaces, most of the bold typefaces are within the upper right quadrant (Figure 73). This indicates that many bold typefaces have similar proportions to typefaces that are experimentally found to improve legibility. However, four bold typefaces are not within the upper right quadrant. Meta Bold, Officina Sans Bold, and Trade Gothic Bold are within the lower right quadrant, and Rotis Sans Bold is within the lower left quadrant.

One typeface, Stone Sans Bold, has a notably larger stroke width than the other bold typefaces (40.0% stroke width) (Figure 72). Based on the scientific review, the legibility of typefaces at this stroke width is not clear. As discussed in Chapter 4 (section 4.3.7), Franklin Gothic Heavy (39% stroke width, 98% letter width) is found to improve legibility for lowercase letters and words (but not uppercase letters) (Sheedy et al., 2005). Stone Sans Bold has a much higher letter width value (108.8%), therefore it is possible that this typeface is more legible than Franklin Gothic Heavy, however this cannot be confirmed based on available data.

6.4 DISCUSSION: IMPLICATIONS FOR INCLUSIVE TYPOGRAPHY GUIDELINES

The lack of standardisation in typeface weight nomenclature is considered in section 6.3.2.1. Figure 70 and Table 18 show empirically that inclusive typography guidelines recommending a range of typeface weights from regular to bold (section 2.4.3.3) are ultimately recommending an enormous range of stroke widths for readers with low vision. Even if guidelines recommend specific weights, this analysis evidences that regular and bold weight names represent a range in parameter values, particularly bold weights. This subject is revisited again in the final paragraph of this section.

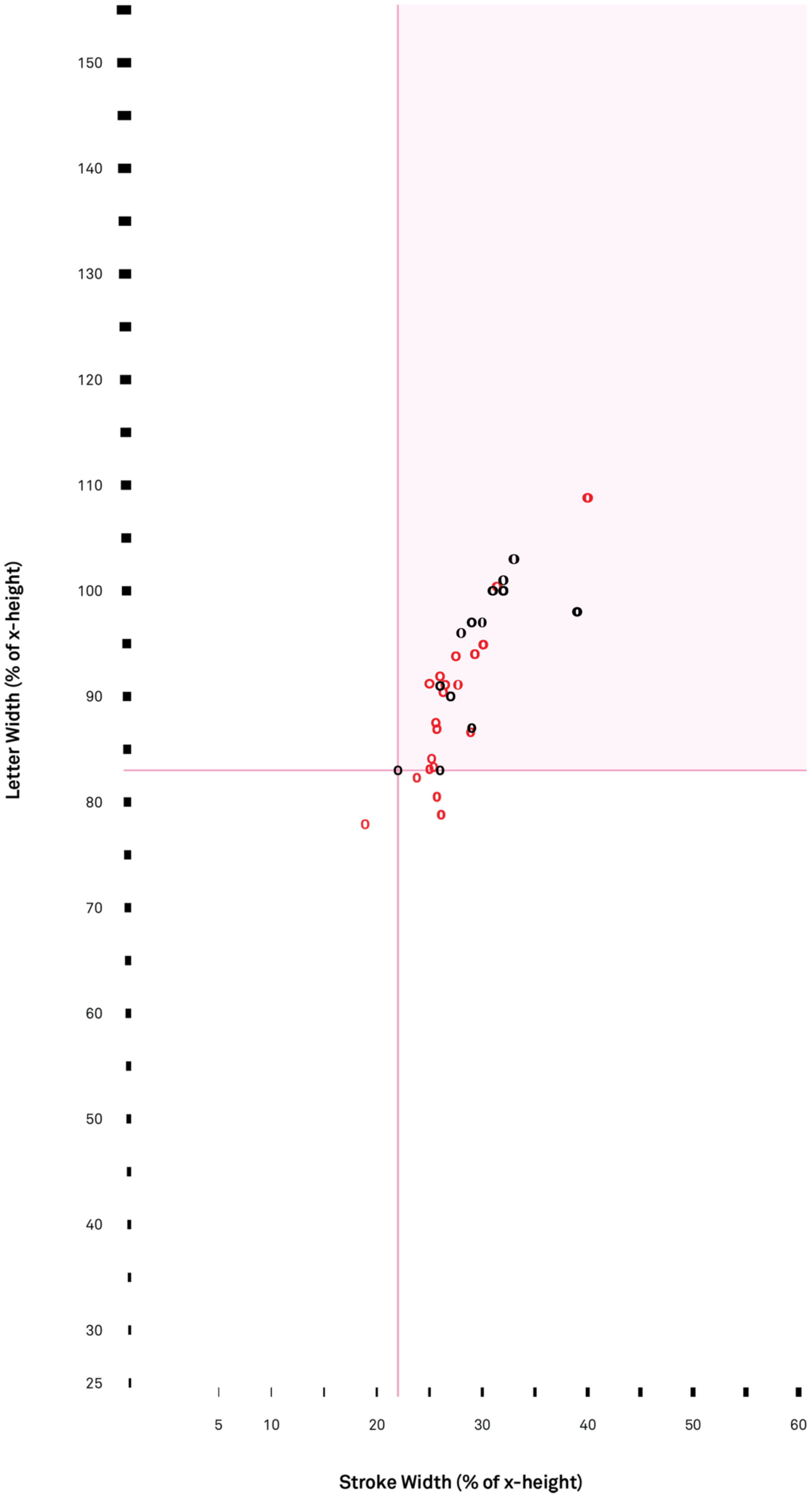


Figure 72: Relationship between typefaces experimentally found to be more legible (black) and bold typefaces (red). Upper right quadrant shaded pink.

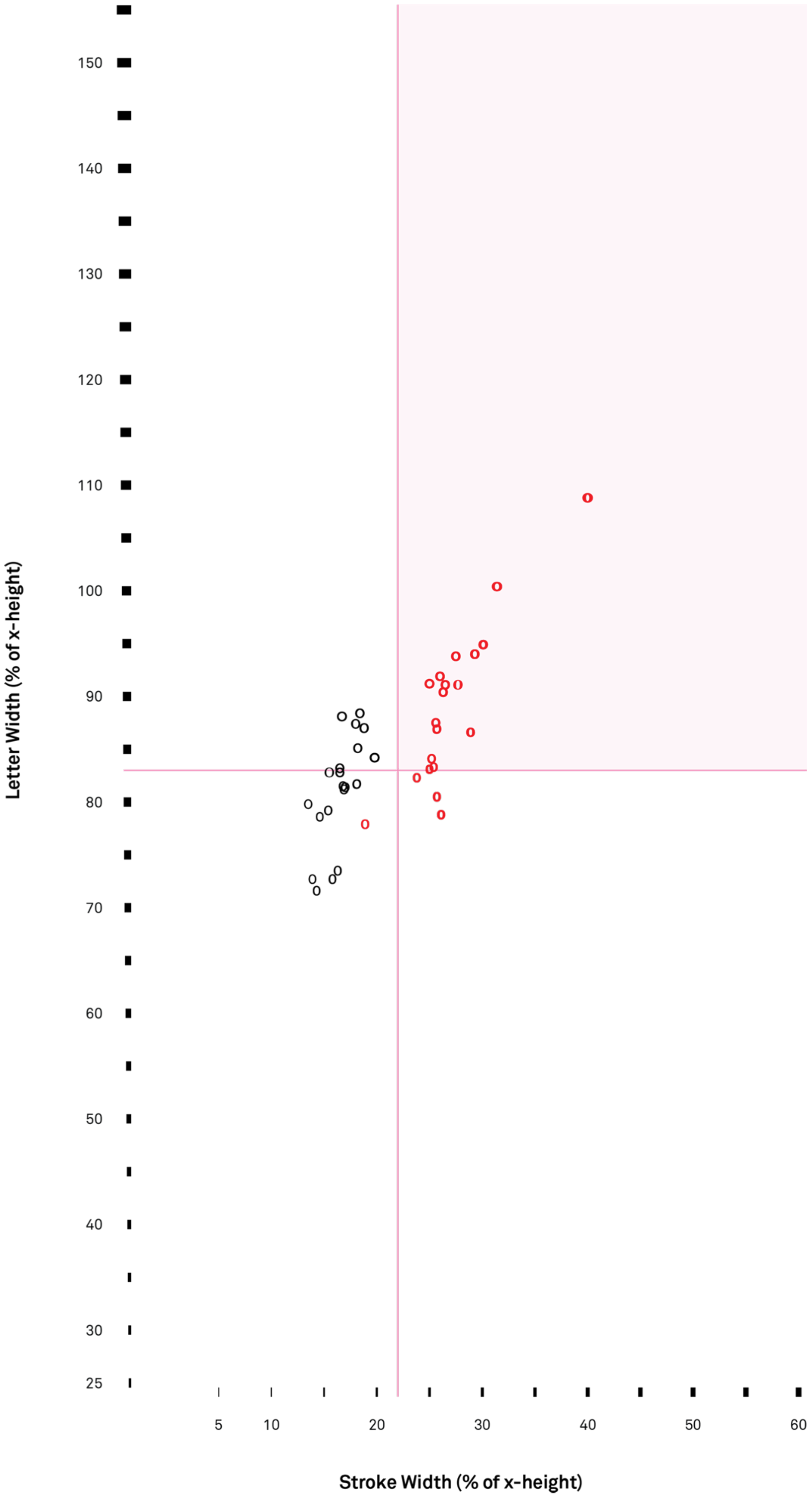


Figure 73: Most bold typefaces (red) are in the upper right quadrant (pink area) where there are no regular typefaces (black).

The scientific review (section 4.4.2) indicates that typefaces ranging from 22-33% stroke width improve reading performance in the context of low vision. As considered in section 6.3.2.2, Figure 71a illustrates that roman regular typefaces have stroke width values below 22%, and do not have similar proportions to typefaces experimentally found to improve legibility (i.e. within the upper right quadrant). This analysis indicates that typefaces commonly used to set continuous text are not optimal for low vision readers.

As discussed in section 6.3.2.2, Figure 72 illustrates that the majority of roman bold typefaces have proportions similar to typefaces found to improve reading performance. However there are exceptions to this, with some bold typefaces having higher and lower stroke width and letter width values. This analysis suggests that many bold typefaces, but not all, could potentially improve legibility for readers with low vision.

Ideally, inclusive print guidelines would recommend typefaces that are experimentally found to improve legibility (e.g. Gill Sans Bold) for the setting of continuous text. However, as can be seen from the scientific review, a limited number of typefaces fall into this category. A reasonable second choice is for inclusive print guidelines to recommend typefaces with similar parameter values to typefaces found to improve legibility. This would still represent an improvement to guidelines which lack specificity (section 2.4.3.3) and an evidentiary basis (section 1.2.4). A third—and less ideal—option is for inclusive print guidelines to recommend roman bold typefaces for continuous text versus for emphasis (e.g. RNIB, 2017). Based on the lack of standardised nomenclature, particularly for bold weights, any such recommendation would lack specificity. However, the analysis (Figure 72) suggests that this would still result in recommending typefaces that are more likely than not to improve legibility as compared to regular weight typefaces. Inclusive typography guidelines and future legibility research are considered further in Chapter 7.

6.5 SUMMARY

This chapter formalises design knowledge on sans serif text typeface proportions, integrates this with scientific knowledge consolidated in Chapter 4, and develops interdisciplinary typeface knowledge. The survey of design sources reveals twenty typeface families appropriate for the generation of design knowledge. Visualisations and statistical analyses of these typefaces indicate that when inclusive typography guidelines recommend a range of weights from regular to bold, they are ultimately recommending a large range of stroke width values for readers with low vision. Visualisations of interdisciplinary knowledge indicate that roman regular typefaces commonly used for setting text are not optimal for low vision readers. Visualisations further suggest that the majority of roman bold typefaces—but not all—may potentially improve reading performance for people with low vision. Three options are proposed for future inclusive print guidelines: (1) Recommend typefaces experimentally found to improve legibility (ideal), (2) Recommend typefaces that share proportions with typefaces found to improve legibility (reasonable), and (3) Recommend roman bold typefaces (less ideal).

Chapter 6 presents the final practice work of this PhD research. Next we turn to Chapter 7 which summarises the scientific, design, and interdisciplinary knowledge on typeface legibility for low vision adults contributed through the PhD investigation.

CHAPTER 7. CONCLUSIONS AND FUTURE RESEARCH

7.1 INTRODUCTION

This final chapter of the PhD thesis summarises the scientific, design, and interdisciplinary knowledge on typeface legibility for low vision adults contributed through this investigation. First, an overview of the PhD research is provided, followed by summaries of the practice chapters and the findings of the investigation. The main contributions to knowledge are then presented in the context of future legibility research and the development of evidence-based inclusive print guidelines. Finally, this chapter considers the scope of the PhD investigation, and points to future directions for inclusive typography research.

7.2 OVERVIEW OF RESEARCH

The intention of this PhD research was to contribute to the development of inclusive typography knowledge on low vision legibility. The investigation was driven by the research question: What is the optimal typeface stroke width for low vision adults? My research addressed this question through a practice-based interdisciplinary approach focused on the quantitative analysis of typefaces.

The research began with a focus on designing a typeface for the controlled investigation of stroke width. It was through my typeface design practice that the methods for the quantitative analysis of typefaces emerged. The PhD then changed direction to focus on developing interdisciplinary typographic knowledge through information visualisation. This entailed consolidating scientific knowledge by measuring and visualising the proportions of typefaces found to improve reading performance. A design phenomenology study was conducted, generating design knowledge by measuring and visualising the proportions of typefaces commonly employed in typographic practice. Finally, through an integration of this scientific and design knowledge, the relationship between the proportions of typefaces found to improve legibility and the typefaces commonly utilised in design practice was determined. This interdisciplinary knowledge on typeface legibility for low vision readers makes an original contribution to knowledge informing future inclusive print guidelines and legibility research.

7.3 PRACTICE CHAPTER SUMMARIES AND FINDINGS

Chapter 4 addressed the research question through a scientific review, analysing ten scientific studies that test the influence of stroke width on legibility. Through the quantitative analysis of typefaces employed as experimental test material, the relationship between typeface parameters and reading performance was elucidated. Based on the data, typeface stroke width values ranging from 22-33% improve reading performance in the context of low vision.

Chapter 5 described the development of a laboratory typeface, which demonstrated that experimental test material can be designed for the controlled study of stroke width, while reflecting design practice. A design-informed laboratory typeface for the controlled study of stroke width is yet to be published in the research literature. Most importantly, this chapter revealed the origins of the quantitative analysis of typefaces, as the laboratory typeface was based upon a consolidation of scientific knowledge and the generation of design knowledge.

Chapter 6 presented the generation of design knowledge and development of interdisciplinary knowledge. Through a survey of multiple design sources, twenty sans serif typefaces were revealed as representative of design practice, and validated to serve as the basis for the generation of design knowledge. The analysis of these typefaces showed that roman regular typefaces range from 13.5-19.8% stroke width, with a median of 16.7%. Roman bold typefaces range from 18.9-40.0% stroke width, with a median of 26.1%. Visualisations of interdisciplinary knowledge evidenced that roman regular typefaces commonly used in design practice have lower stroke width values than typefaces found experimentally to improve reading performance in the context of low vision. In contrast, many roman bold typefaces have proportions similar to typefaces found to improve reading performance. These visualisations addressed the research question in the context of design practice, clarifying optimal typeface weights for low vision adults. The above findings result in an original contribution of knowledge in the emerging area of inclusive typography, discussed in the following section.

7.4 CONTRIBUTION TO KNOWLEDGE

7.4.1 SUMMARY OF CONTRIBUTION

This PhD thesis develops a foundation of interdisciplinary—scientific and design—knowledge on typeface legibility for low vision adults. Through a quantitative analysis of typefaces, involving the measurement and visualisation of typeface proportions, this research elucidates the relationship between typeface parameter values and reading performance. This research contributes to future legibility research and the development of evidence-based inclusive print guidelines.

My three main contributions to knowledge in the emerging field of inclusive typography are as follows:

1) Visualisations of scientific knowledge: Stroke width and letter width values of typefaces found to have higher and lower legibility in the context of low vision reading;
2) Visualisations of design knowledge: Stroke width and letter width values of sans serif text typefaces utilised in design practice;
3) Visualisations of interdisciplinary knowledge: Relationship between stroke width and letter width values of typefaces found to have higher legibility (scientific knowledge) and typefaces utilised in design practice (design knowledge).

7.4.2 SCIENTIFIC KNOWLEDGE

Visualisations of scientific knowledge make an original contribution to the understanding of optimal stroke width values for low vision reading. My analysis established that that the optimum stroke width for legibility in a low vision context is not a regular weight typeface and is found at stroke width values above this. How much above is not clear, however based on the analysis of scientific studies, typefaces with stroke width values of 22-33% improve reading performance as compared to regular weight typefaces (section 4.4.2). Future legibility research would benefit from examining stroke widths of both intermediate (i.e. between 22-29%) and higher (i.e. above 30%) values in order to determine an optimum for low vision adults.

My analysis demonstrated that bolder typefaces are more legible than regular weight typefaces, however the isolated influence of stroke width, letter width, and letter spacing remains unclear. Visualisations illustrated that bolder typefaces generally have larger letter widths (e.g. Figure 68), and it is likely that their improved legibility is due to both parameters. My analysis showed that typefaces found to improve legibility all had letter widths above 83%. My analysis of Luckiesh and Moss (1940) also illustrated that an increased stroke width can improve legibility even if letter width is decreased (section 4.3.3). My analysis of Roethlein (1912) demonstrated empirically that a bolder typeface with a smaller letter width (i.e. Condensed Bold) may not be more legible than a regular typeface (section 4.3.2).

While legibility studies which test commercial typefaces cannot assess the isolated influence of parameters, my research has illustrated that a quantitative analysis of these typefaces can shed light on the relationship between parameter values and reading performance. This is the first review of this kind; an analysis of experimental test material of ten scientific studies with the purpose of clarifying the relationship between typeface stroke width and letter width values, and reading performance. While the research literature contains examples of the measurement of typeface parameters with the purpose of relating them to reading performance (e.g. Xiong et. al., 2018), this approach is not common (section 3.4.5). My research helps to establish the utility of a quantitative analysis of typefaces within legibility research studies. While my research is useful to legibility researchers focused on stroke width, the detailed explanation of all methods within this thesis offers a methodology to future researchers who may be interested in undertaking similar quantitative reviews.

7.4.3 DESIGN KNOWLEDGE

Visualisations of design knowledge contribute to the development of future experimental test material that reflects design practice. As argued in Chapter 3 (section 3.2.1), design knowledge must be incorporated into scientific legibility research in order for it to be applicable to design practice. Yet, how can legibility researchers be expected to appropriately reflect design practice when there is a lack of formalised typeface design knowledge? When I set out to design my laboratory typeface, there was little available information that I could base its parameters on. This led to the generation of design knowledge; collecting data on text typeface stroke width and letter width values. My research contributes the following information, intended to be useful for future legibility researchers: (1) Twenty sans serif typefaces that are representative of typography practice (section 6.2.2), and (2) An analysis of stroke width and letter width values of sans serif text typefaces (section 6.3.2.1).

Visualisations of design knowledge also contribute to the development of evidence-based inclusive print guidelines. Visualisations of design knowledge (e.g. Figure 70), supported by summary statistics (Table 18), confirmed empirically that typeface weight nomenclature represents a range of numerical values. This validates my assessment of inclusive typography guidelines as lacking specificity (e.g. section 2.4.3.3). For example, guidelines that recommend regular and bold weight typefaces (e.g. RNIB & ISTD, 2007) are ultimately recommending an enormous range of stroke widths from 13.5-40.0% for low vision readers (see section 6.3.2.1). Even a recommendation of regular or bold weights results in recommendations of stoke width ranges of 6.3% and 21.1%, respectively. The range is particularly high in bold typefaces; the largest measured stroke width was more than double the value of the smallest. While it is established within the design community that typeface weight nomenclature does not represent numerical stroke width values (e.g. Bigelow, 2019), my analysis demonstrated this empirically. This is the first published quantitative analysis of the relationship between typeface weight nomenclature and numerical values. My analysis indicates that inclusive print guidelines would be improved if specific stroke width values (e.g. 22-33%) and/or specific typefaces (e.g. Franklin Gothic Medium, Gill Sans Bold) are recommended for low vision readers, versus typeface weights (e.g. regular, bold).

7.4.4 INTERDISCIPLINARY KNOWLEDGE

Visualisations of interdisciplinary knowledge make an original contribution to the development of evidence-based inclusive print guidelines. Figure 71a established that regular weight typefaces commonly used to set continuous text have stroke widths lower than typefaces experimentally found to improve reading performance. This analysis indicated that regular weight typefaces are not optimal for low vision legibility (section 6.4.2). Based on this analysis, inclusive print guidelines would be improved by reconsidering the recommendation of regular weight typefaces. Figure 72 showed that many bold typefaces share proportions with typefaces experimentally found to improve legibility. This analysis indicated that many bold typefaces, but not all, may improve legibility for readers with low vision.

In Chapter 6 (section 6.4.2) I proposed three options for inclusive print guidelines. First and ideally, guidelines can recommend typefaces that are experimentally found to improve legibility, for the setting of continuous text. The scientific review in Chapter 4 contributed an analysis of these typefaces (Table 1), which can be practically applied by designers in their choice of typefaces for a low vision audience. A reasonable second option is for guidelines to recommend typefaces with similar proportions to typefaces found to improve reading performance (see section 6.3.2.2). This recommendation has practical utility because only a limited number of typefaces have been tested experimentally. It is not possible to confirm that a typeface with similar stroke width and letter width values to legible typefaces also performs well. However, recommending such typefaces still represents an improvement to guidelines which lack specificity (section 2.4.3.3). The final and least ideal option for inclusive print guidelines is to recommend bold typefaces for continuous text versus for emphasis (e.g. RNIB, 2017). The analysis in Chapter 6 (section 6.3.2.2) demonstrated that most bold typefaces share stroke width and letter width proportions with typefaces found to improve legibility in the context of low vision. However, not all bold typefaces share these proportions and may have stroke widths or letter widths that fall outside of optimal ranges. While not ideal, a recommendation of bold typefaces again represents an improvement to guidelines which recommend regular weight typefaces (RNIB & ISTD, 2007) and bold “sparingly” for emphasis (RNIB, 2017, p.1) (section 2.4.3.3). These three recommendations offer practical contributions to the development of evidence-based inclusive typography guidelines.

As introduced in Chapter 2 (section 2.4.2), inclusive print guidelines are defined primarily by size recommendations (RNIB & ISTD, 2007). UKAAF (2019) recommends 12 and ideally 14 point for Clear Print documents, which are designed to reach wider audiences. UKAAF (2019) recommends 16 and ideally 18 point for the more specialised Large Print format (Figure 7). As demonstrated in Chapter 2, guidance on the choice of typefaces themselves is broad. This research seeks to inform print design guidelines with a higher degree of specificity, and thereby increase the proportion of the population able to access text. I have outlined recommendations above, and here consider their practical implementation.

While my analysis indicates that stroke width values ranging from 22-33% x-height improve legibility in the context of low vision, this would be difficult for many people using guidelines to implement. The guidelines are for “anyone producing clear print or large print documents” (UKAAF, 2019) and not necessarily trained designers. While specifying a stroke width range is useful, a more practical guideline would offer a list of typefaces that have either been found to improve legibility experimentally or share proportions with them. Based on my research, sans serif typefaces that have been shown to improve legibility and are commonly available include: Gill Sans Bold, Helvetica Bold, Arial Bold, Verdana Bold, Franklin Gothic Medium, and Franklin Gothic Demi (see Chapter 4, Appendix 2).

If bolder typefaces are recommended within guidelines as my analysis suggests, how does this influence size recommendations? I am not advocating decreasing the recommended sizes for specific formats; 12-14 point for Clear Print and 16-18 point for Large Print. My analysis does suggest however that at each given type size, a bolder typeface should increase the number of people able to read that text. This is important because vision rehabilitation works according to the principle that text should be magnified as little as possible to reach a person’s CPS. If a person with low vision is able to read bolder text at a smaller size, this could increase the number of letters that fit into the field of view. As discussed in section 2.2.3, the number of letters that can be seen at a time impacts reading performance (Legge, 2007).

There remains significant work to be done in the development of evidence-based inclusive typography guidelines. However based on my analyses, I propose that recommending weights above ‘regular’ is an improvement. While research is lacking in the area of typeface legibility, there is enough evidence, in my opinion, to suggest that recommending typefaces with increased weight would increase the percentage of the population able to access text. As always, more research needs to be done.

7.5 LIMITATIONS, INSIGHTS, AND FUTURE RESEARCH

7.5.1 INTRODUCTION

The following sections consider the scope of the PhD investigation, with a view toward future research. First, I discuss letter width and letter spacing; two parameters which were of secondary focus in my thesis and are worthy of further examination. Second, I consider research which suggests that the influence of typeface characteristics on legibility changes depending on the underlying cause of visual impairment. Third, and finally, I revisit the topic of commercial versus laboratory typefaces as experimental test material, and my fundamental change in perspective through the PhD research process.

7.5.2 TYPEFACE PARAMETERS

This PhD investigation was primarily focused on stroke width. As reviewed in Chapter 2 (section 2.4), numerous typeface parameters influence legibility in the context of low vision: character size, presence or absence of serif, stroke width, letter width, monospaced versus proportional typefaces, and letter spacing. Letter width and letter spacing stand out as particularly important typeface parameters worthy of further investigation.

As addressed in section 7.4.2, the increased legibility of bolder typefaces is likely due, in part, to increased letter width. As introduced in Chapter 2 (section 2.4.3.4), size is a critical typographic variable for low vision readers, and evidence suggests that a horizontal increase in size (i.e. increased letter width) has a similar legibility benefit (Arditi, 2004). My analysis indicated that typefaces found to have increased legibility all had letter widths above 83% (section 6.3.2.2). Further research would be useful to clarify the relationship between letter width and legibility, and inform inclusive print recommendations which generally do not address this parameter.

Letter spacing is another fruitful area for future legibility research. As reviewed in Chapter 2 (section 2.4.3.6), research suggests that letter spacing above the standard spacing of a typeface does not improve legibility, based on an experiment testing Courier (Chung, 2012). However, the study by Beveratou (2016) finds a legibility benefit for increased letter spacing and leading of a proportionally spaced sans serif typeface. Further research would be useful to test the legibility of proportionally spaced typefaces with increased letter spacing, which could be particularly beneficial for bold typefaces which often have decreased letter spacing.

My analysis suggested that the increased legibility of Courier Bold and Maxular Rx Bold is likely due to a combination of increased letter width and letter spacing (section 4.3.9). The study by Xiong et al. (2018) also examines spacing—a measure which included letter width and letter spacing—finding that typefaces with larger spacing permit smaller reading acuity and CPS (section 4.3.9). Are monospaced typefaces more legible (Mansfield et al., 1996) because they have wider letters overall and larger letter spacing? Could proportionally spaced typefaces with increased letter widths and letter spacing achieve equivalent legibility? Aspects of Maxular Rx address this, as it is a proportionally spaced typeface with larger letter widths and letter spacing. Maxular Rx Bold permits smaller reading acuity and CPS than Helvetica and Times, however does not perform better than Courier (Xiong et al., 2018) (section 2.5.3). Investigating proportionally spaced typefaces with larger letter width and letter spacing would be valuable, as continuous text is typically set in proportionally spaced typefaces.

7.5.3 UNDERLYING CAUSE OF VISUAL IMPAIRMENT

This thesis has only touched on evidence suggesting that typeface characteristics may influence reading performance differently depending on the underlying reasons for visual impairment. Arditi’s (2004) study using prototype software (Font Tailor) finds that low vision participants choose different parameter values in order to customise typefaces to meet their own visual needs. In the context of customising typefaces, Beveratou (2016) notes that individuals with different types of visual impairment have different legibility needs. Shaw's (1969) study demonstrates empirically that typographic variables influence legibility differently, depending on the underlying reason for visual impairment. Shaw states:

> "Readers with cataract were not helped by increase in size so much as increased weight of print … The group of readers with glaucoma were affected more than other groups by typographic changes, and size and weight were both important for them. Readers with macular degeneration were helped by increases in print size and a change to a sans serif type but, surprisingly, not by an increase in weight." (Shaw, 1969, p.62).

Shaw (1969, p.62) concludes that "despite these detailed differences, the results do not present any real conflict in practical terms. In no instance was a typographic factor helpful to one group of readers, but positively bad for one of the other groups." My typeface recommendations (section 7.4.4) are based on the assumption that increased boldness does not result in decreased legibility for any one user group. However, further studies are needed to confirm this. It should also be noted that of the ten studies analysed in Chapter 4, only four test participants with visual impairments (Shaw, 1969; Mansfield et al., 1996; Beveratou, 2016; Xiong et al., 2018). While the argument can be made that experiments testing normally sighted participants at threshold sizes can be applied to people with low vision who often read at their acuity limit, Shaw's (1969) research is a reminder that different visual abilities influence the results of legibility studies. The area of inclusive typography needs not only more experimental research, but research with low vision participants which address the question of how different underlying conditions influence typographic needs.

7.5.4 COMMERCIAL VERSUS LABORATORY TYPEFACES REVISITED

In one important area, my perspective has fundamentally changed through the PhD research process. I began this investigation passionately focused on the lack of controlled investigation in legibility research studies. Responding to the predominance of studies employing commercial typefaces as experimental test material, my initial focus was on designing a laboratory typeface for the controlled study of stroke width. Interestingly, my final contributions to knowledge revolve around the measurement of commercial typefaces, and the analysis of scientific studies which test these.

Does it really matter that we have not disentangled the isolated influence of stroke width and letter width on the legibility of bold typefaces? The point is that bold typefaces are generally more legible than regular typefaces. The studies that were analysed in Chapter 4 do not control typeface parameters, and in this way employ test material that is directly applicable to typefaces utilised by designers, which also simultaneously vary in stroke width and letter width. It should be noted that while two studies I analysed test a laboratory (Beier & Oderkerk, 2019a) and a propriety typeface (Beier & Oderkerk, 2019b), parameters were not controlled which is a similar approach to testing commercial typefaces.

Reflecting on my typeface design work, I made many compromises to control parameters (section 5.4); each one taking the typeface one step further from the design of commercial typefaces. Laboratory typefaces developed for the controlled study of parameters are generally more theoretical than studies which directly test the legibility of commercial typefaces in use by designers. Controlled studies contribute to the theoretical understanding of the relationship between typeface parameters and legibility. However in the case of stroke width, we have an understanding of this parameter; stroke width influences legibility, and the optimum lies at intermediate values (Arditi et al., 1995a). If we are interested in what stroke widths are optimal for low vision reading in the context of typefaces that designers use, we must test the legibility of commercial typefaces. The ‘why’ does not matter as much as ‘which’ typefaces are potentially more legible.

In the final stages of thesis writing, when I began to make recommendations, I realised that ideally guidelines would advise the use of typefaces that have been experimentally tested, like Gill Sans Bold by Shaw (1969). These recommendations depend on studies that test commercial typefaces. It must be noted however, that the utility of these studies relies on publishing images of their experimental test material. Scientific studies testing commercial typefaces then have the further benefit of being replicable, and offer an opportunity for other researchers to undertake subsequent analyses, as I have done in Chapter 4’s scientific review.

Testing commercial typefaces is an extremely valuable method for legibility research in my opinion. This represents a reversal of my perspective at the beginning of this research. At this point, my recommendation is for legibility researchers to test the most commonly used typefaces. For example, a study could test a range of weights of the most popular typefaces (e.g. Helvetica Regular, Demi, Bold, etc.) with low vision participants. A study by Burmistrov et al. (2016) employs such test material, examining four weights of Helvetica Neue (Ultra-Light, Light, Normal, Bold) in Cyrillic. Experimental studies utilising test material like this are more directly applicable to typography practice.

What is the optimal typeface stroke width for low vision adults? My initial approach to this research question was focused on determining the isolated influence of stroke width on legibility. I now see this as a theoretical question. My PhD research has ended up addressing this question in a practical manner, examining the stroke widths of commercial typefaces that have been experimentally found to be more legible, and commercial typefaces that have similar proportions.

In the end, I am surprised to hinge my knowledge generation on these studies testing commercial typefaces, that I had been so quick to criticise early in my design research career. While experimental test material can often be criticised, my research highlights many scientific studies that contribute to typographic knowledge construction. I am grateful that scientists including Roethlein, Arditi, Sheedy, and Legge have taken on these questions so pivotal to design practice. Design researchers are now in a position to collaborate with scientists or take on these questions themselves, employing psychophysical methods. Ultimately, my research practice not only evidences the value of design approaches to typeface legibility, it honours the value of scientific research.

APPENDIX

Appendix 1. Include 2009 conference paper (von Ompteda, 2009).

Innovation in Inclusive Typography: A Role for Design Research

Karin von Ompteda, Communication Art & Design, Royal College of Art
info@inclusivetype.org

Abstract

As the world's populations age, it is increasingly critical that designers produce accessible communications for people with low vision. Currently, inclusive typography guidelines are vague and do not yet rest upon a strong scientific foundation. As such, there is insufficient knowledge to inform the design of typefaces for readers with low vision. Legibility research is largely conducted from within the scientific community, and the challenge of directly applying this knowledge to the practice of typeface design will be discussed. This paper will conclude that design research—acting as an intermediary between scientific research and design practice—has a potentially exciting contribution to make toward innovation in people centered design.

Key words

Inclusive Design, Typeface Design, Legibility, Practice-Led Design Research

Introduction

This paper is motivated by the unprecedented aging of the world's populations (United Nations, 2002) and the associated increased prevalence of age-related eye disease (WHO, 2004). The majority of people with visual impairments are not blind, having significant remaining vision known as low vision (Arditi, 2004). Low vision compromises sight through blurring, patchiness and loss of central or peripheral vision, and in 2002 was estimated to affect approximately 124 million people worldwide (WHO, 2004).

In the United Kingdom, approximately two million people have significant sight loss, the majority of which are 65 years of age and over (RNIB, 2008). Macular degeneration, glaucoma, cataract and diabetic retinopathy are leading causes of visual impairment; the risk of each increasing with age. Population aging is more rapid in developed nations, with one in five people currently aged 60 years or over in the UK, and a projected ratio of one in three by 2050 (United Nations, 2007). The number of people in Britain with low vision will rise dramatically over the coming decades (RNIB, 2005).

Difficulty with reading is a vital concern for people living with low vision (Legge, 2007), and a central goal of vision rehabilitation is improving access to written materials (Arditi, 1996). Low vision can be defined functionally as a visual impairment resulting in the inability to read the newspaper at a standard distance (40 cm) with best optical correction (Legge, 2007). Access to text is fundamental to participation in modern society, and the design community has an increasingly critical mitigating role to play in this context through the production of inclusive typographic design.

Inclusive design is a response to the diverse demands of today's consumers, especially those who are elderly or disabled (Clarkson et al., 2003). This approach to architectural, product and communication design, seeks to meet the needs of the largest user-group possible, whilst taking into consideration the goals of commerce (Clarkson et al., 2003). In the context of communication design, inclusive typographic principles are central, providing tools which fundamentally influence the number of individuals able to continue reading with low vision. Visual impairment organizations like the Royal National Institute of the Blind in the United Kingdom and Lighthouse International in the United States are primary sources for guidelines on how to design for readers with low vision (e.g. RNIB, 2006; Arditi, 2009). Inclusive typographic principles are increasingly being adopted by practicing designers, yet an examination of the underlying literature suggests that the research often does not translate into specific guidelines.

A recent review of typography for readers with low vision in the *Journal of Visual Impairment and Blindness* concluded that "research has not produced consistent findings and thus that there is a need to develop standards and guidelines that are informed by evidence" (Russell-Minda et al., 2007). Another review discussed the "little scientific typographical research directly related to this user group" (Perera, 2001). A recent empirical paper in *Ophthalmic and Physiological Optics* examining typography for readers with mild to moderate vision loss stated that "the scientific basis for the guidelines is elusive at best" (Rubin et al., 2006). One set of typography guidelines for readers with impaired vision refers to "little reliable information on the comparative legibility of typefaces …" (Arditi, 2009). Typographic guidelines can also exist within legislation like the Americans with Disabilities Act; an aspect of which was referred to as "written without sufficient research to specify such ranges with confidence" (Arditi, 1996). Legibility research has thus far not been able to inform specific evidence-based inclusive typographic design principles.

Typeface legibility research

Inclusive typography guidelines are based largely on scientific legibility research; the term 'legibility' referring to the perceptual properties of text that influence readability (Legge, 2007). Legibility research is often conducted using what are called 'psychophysical' methods. Psychophysics is the study of the relationship between physical stimuli and perceptual responses (Norton et al., 2002); in this case, the relationship between the physical properties of text (e.g. character size, typeface) and reading performance (Legge, 2007). The psychophysical study of reading often employs two legibility metrics: reading acuity and reading speed. Reading acuity is the minimum size, or equivalently the greatest distance, at which text can be read, and is an appropriate legibility metric in the context of low vision because these people often read at their acuity limit (Arditi, 1996). Reading speed is another common legibility metric, preferred by some researchers because it more naturally resembles an ordinary reading situation. It should be stated that legibility is not an invariant property of text, as it is ultimately determined by a reader's visual processing capabilities (Legge, 2007). As such, good reading performance is achieved through congruence between the physical properties of text and the visual processing abilities of the reader (Legge, 2007).

What are the properties of text that influence the reading performance of people with low vision? Type size is a critical typographic variable for low vision reading (Legge, 2007) and inclusive typography guidelines are defined primarily by larger type size recommendations (RNIB, 2005). Yet due to economic and aesthetic reasons, text is often not presented at sufficiently large sizes. In the context of print for example, increased size translates into larger and more costly documents (RNIB, 2005). Further, many designers are thought to prefer small type for aesthetic reasons (Manuelli, 2004), and the cultural connotations of large type with old age can deter both designers' clients and end-users alike (Ivinski, 2000).

Thus, people with low vision who typically have low acuity, often must read at their acuity limit (i.e. smallest readable size). Research has shown that at these ‘threshold’ type sizes, typeface design has an influence on reading performance (Morris et al., 2002).

While evidence supports the important role that typeface design plays in low vision reading (Shaw, 1969; Mansfield et al., 1996; Russell-Minda et al., 2007), the influence of specific parameters is not well understood (Arditi & Cho, 2005). For example, experiments have suggested that typeface weight or ‘boldness’ influences low vision reading (Shaw, 1969; Arditi, 2004), yet optimum boldness values remain unknown. A range of values are currently recommended by inclusive typography guidelines (e.g. RNIB, 2005) and two print typefaces designed for low vision readers—APHont (Kitchel, 2004) and Tiresias LPFont (Perera, 2001)—differ in this parameter. It should further be noted that, as of yet, there exists no experimental evidence to support the claim that these typefaces improve reading performance for people with low vision. This is congruent with the notion that there is insufficient knowledge to inform specific inclusive typography principles or typeface design for people with low vision. This paucity of knowledge can be attributed to a lack of research within the area of typeface legibility for low vision, as well as research methodologies which have left questions unanswered.

Inclusive typography: why a lack of knowledge?

Typeface legibility has not been a major research interest over the past four decades (Legge, 2007). Some have gone so far as to reference “the current ‘low status’ of legibility research and its demise as a ‘research program’…” (Lund, 1999). This lack of research has been attributed to the practical challenges of manipulating typefaces for experimental purposes, as well as the perception that a creative design artifact is not amenable to “quantitative stimulus description” (Legge, 2007). Of existing typeface legibility research, low vision readers have historically not been a focus (Russell-Minda et al., 2007). Furthermore, methods of testing have left questions unanswered.

Scientists investigating typeface legibility are known to commonly test commercially available typefaces which differ in numerous parameters (e.g. Mansfield et al., 1996). While such experiments can tell us that one typeface is more legible than another, they cannot with any certainty tell us why. Thus, while evidence supports the important role that typeface design plays in low vision reading (Russell-Minda et al., 2007), the influence of specific parameters has remained elusive. This paucity of controlled investigation within the field of legibility research has been the basis for criticism by designers (Dyson, 1999; Lund, 1999) as well as scientists (Arditi & Cho, 2005), and has further discouraged research interest in the area (see above). Yet only with relatively recent digital technology has it become feasible to design experimental typefaces whose parameters can be adjusted systematically to assess their influence on low vision reading (Arditi, 1996). While the research community has discussed the value of such a methodology (Russell-Minda et al., 2007) experimental literature in the area is still rare.

The work of vision scientist Aries Arditi (and colleagues) has made a notable contribution to this methodology, having taken advantage of the opportunities of digital technology. Arditi has investigated the influence of numerous typeface parameters on legibility, through systematic and carefully controlled study made possible through his custom laboratory typefaces (e.g. Arditi et al., 1995a; Arditi et al., 1995b; Arditi, 2004; Arditi & Cho, 2005). While Arditi has read a significant amount on the subject of type design, as well as involving a typeface designer in some of his research projects, he has not asked a professional typeface designer to develop his experimental typefaces for him and took on this task himself. In the context of contemporary typefaces designed for continuous reading, his are unconventional in their construction and illustrate a distance between scientific methods and design practice. While Arditi himself acknowledges freely that his fonts are not aesthetically pleasing, he asserts that they are specifically designed to answer particular questions about the impact of typographic variables on legibility (personal communication, August 12, 2008).

When typeface parameters are abstracted in the service of experimental control, does this come at the cost of applying the research to design practice? In an investigation employing a custom typeface to test the influence of boldness on legibility (Arditi et al., 1995a), the white internal spaces of many of the letters in the boldest font closed in with ink (e.g. think of the two white spaces within the capital 'B') resulting in distortions of letterforms. This would not occur within a professional commercial typeface, as bolder fonts are designed with decreased weight of internal strokes, for example, in order to maintain the white spaces necessary for letter recognition (see: M vs. **M**). In this experiment, Arditi held stroke weight constant within letters (i.e. a monoline font), which has little effect in the lighter weight fonts, but made it unfeasible for the boldest font to resemble conventional type.

While typographers (e.g. Carter et al., 2007) would agree with Arditi's results—magnitude of boldness will eventually decrease legibility (Arditi et al., 1995a)—this experiment does not inform designers of the boldness value where this legibility decrease occurs. This leads one to ask: should experimental typefaces resemble commercial ones? According to Arditi, if one's goal is to investigate the effects of typographic variables, then this is unnecessary; yet if one's goal is to investigate the effects of variables within the confines of conventional typefaces, then most definitely (personal communication, August 12, 2008). If one were to do the latter, one would have to accept that in the bolder fonts, two parameters would be changing simultaneously: boldness and variation in stroke weight. However, the benefit of such an approach would be in maintaining letterform familiarity across the fonts. There is reason to believe that experimental control and typographic reality do not have to be at odds. One legibility study, co-authored by a typeface designer, developed experimental fonts by modifying the Lucida typeface family (Bigelow and Holmes Inc., San Jose, CA), based on the assumption that one component of legibility is familiarity (Morris et al., 2002).

Conclusions and further research

Scientific legibility research has provided the design community with a wealth of knowledge regarding the role of typography in influencing reading performance. Yet, if we are to understand the specifics which will influence inclusive typographic practice and the development of typefaces for people with low vision, I suggest there is a role for design researchers. Design academics have advocated for legibility researchers to increase their knowledge of typographic practice (Dyson, 1999; Lund, 1999), yet the inception of practice-led design research degrees has created a context within which designers can undertake these investigations underpinning our discipline. Responding to this opportunity, typeface legibility research projects have been initiated at both the Royal College of Art (United Kingdom) as well as the University College of the Province of Limburg (Belgium). This is a new context for legibility research; academic design researchers developing typefaces as test material and running experiments which employ psychophysical testing methods (supervised by scientists). There are two major reasons to believe that design academics have the ability to contribute to this research area. First, typeface designers are well versed in the sophisticated manipulation of typefaces; essentially their test material. Second, these researchers are studying legibility in the context of conventional type, with direct application to design practice and design artifacts as a central goal. These research projects are currently in progress, and it will be exciting to see to what degree interdisciplinary design research can lead toward innovation in inclusive typographic design.

Acknowledgements

This research is funded by the Commonwealth Scholarship Commission in the United Kingdom.

References

Arditi, A (2009). Making Text Legible: Designing for People with Partial Sight [Homepage of Lighthouse International], [Online]. Available: http://www.lighthouse.org/accessibility/legible/ [05 Feb., 2009].

Arditi, A (2004). Adjustable Typography: An Approach to Enhancing Low Vision Text Accessibility. Ergonomics, vol 47, no 5, 469-482.

Arditi, A (1996). Typography, Print Legibility and Low Vision. In: Rosenthal, B and Cole, R Eds, Remediation and Management of Low Vision. St. Louis: Mosby.

Arditi, A, Cagenello, R and Jacobs, B (1995a). Letter Strokewidth, Spacing, and Legibility. Vision Science and its Applications. Washington D.C.: Optical Society Of America.

Arditi, A, Cagenello, R and Jacobs, B (1995b). Effects of Aspect Ratio and Spacing on Legibility of Small Letters. Investigative Ophthalmology, vol 36, no 4, 36(4), 671.

Arditi, A and Cho, J (2005). Serifs and Font Legibility. Vision Research, vol 45, no 23, 2926-2933.

Barratt, M and Walker, S (2002). Typography for Visually Impaired Users [Homepage of Text Matters], [Online]. Available: http://www.textmatters.com/our_interests/guidelines/typog_visual_impaired/ [02 Jan., 2006].

Carter, R, Day, B and Meggs, P (2007). Typographic Design: Form and Communication. 4th Edition. New Jersey: John Wiley and Sons.

Clarkson, J, Coleman, R, Keates, S and Lebbo, C, eds (2003). Inclusive Design: Design for the Whole Population. London: Springer.

Dyson, M (1999). Typography Through the Eyes of a Psychologist. Hyøen, vol 1, 5-13.

Ivinski, P (2000). Type For Sore Eyes. Print, vol 54, no 2, 8.

Kitchel, E (2004). APHont™: A Font for Low Vision [Homepage of American Printing House for the Blind], [Online]. Available: http://www.aph.org/products/aphont.html [01 Oct., 2005].

Legge, G (2007). Psychophysics of Reading in Normal and Low Vision. London: Lawrence Erlbaum Associates.

Lund, O (1999). Knowledge Construction in Typography: The Case of Legibility Research and the Legibility of Sans Serif Typefaces, Thesis (PhD), The University of Reading.

Mansfield, S, Legge, G and Bane, M (1996). Psychophysics of Reading XV: Font Effects in Normal and Low Vision. Investigative Ophthalmology and Visual Science, vol 37, no 8, 1492-1501.

Manuelli, S (2004). Make A Point. Design Week, vol 19, no 9, 16-17.

Morris, R, Aquilante, K, Yager, D and Bigelow, C (2002). Serifs Slow RSVP Reading at Very Small Sizes, but Don't Matter at Larger Sizes. Society for Information Display International Symposium Digest of Technical Papers, vol 33, 1-4.

Norton, T, Corliss, D and Bailey, J (2002). The Psychophysical Measurement of Visual Function. Woburn: Elsevier Science.

Perera, S (2001). LPfont: An Investigation into the Legibility of Large Print Typefaces [Homepage of Tireseas], [Online]. Available: http://www.tiresias.org/publications/reports/lpfont%20report/index.htm [01 October, 2005].

RNIB (2008). Statistics: Numbers of People with Sight Problems by Age Group in the UK [Homepage of Royal National Institute of the Blind], [Online]. Available: http://www.rnib.org.uk/xpedio/groups/public/documents/PublicWebsite/public_researchstats.hcsp [14 Apr., 2008].

RNIB (2006). See it Right: Making Information Accessible for People with Sight Problems. London: Royal National Institute of the Blind.

RNIB (2005). Inclusive Design: Clear and Large Print Best Practice Guide for Designers. Taunton: International Society of Typographic Designers.

Rubin, G, Feely, M, Perera, S, Ekstrom, K and Williamson, E (2006). The Effect of Font and Line Width on Reading Speed in People with Mild to Moderate Vision Loss. Ophthalmic and Physiological Optics, vol 26, no 6, 545-554.

Russell-Minda, E, Jutai, J, Strong, J, Campbell, K, Gold, D, Pretty, L and Wilmot, L (2007). The Legibility of Typefaces for Readers with Low Vision: A Research Review. Journal of Visual Impairment and Blindness, vol 101, no 7, 402-415.

Shaw, A (1969). Print For Partial Sight: A Report To The Library Association Sub-Committee On Books For Readers With Defective Sight. London: Library Association.

United Nations (2007). World Population Ageing 2007. New York: United Nations Publications.

United Nations (2002). World Population Ageing: 1950-2050 [Homepage Of United Nations], [Online]. Available: http://www.un.org/esa/population/publications/worldageing19502050/index.htm [16 Nov., 2007].

World Health Organization (2004). Magnitude and Causes of Visual Impairment. Available: http://www.who.int/mediacentre/factsheets/fs282/en/ [01 Oct., 2005].

Appendix 2. Scientific review parameter data. The table presents studies included in Chapter 4's scientific review, typefaces tested experimentally within each, and their parameter values based on my measurements. This data is reproduced and presented separately for each scientific study in Chapter 4 (Tables 2-12).

Scientific Study	**Typeface Tested**	**Parameter Values (% x-height)**		
		Stroke Width	**Letter Width**	**Letter Spacing**
Roethlein (1912)	Century Old Style	18	75	40
	Century Old Style Bold	28	96	35
	Cheltenham Old Style	21	82	35
	Cheltenham Old Style Bold	31	100	31
	Cheltenham Condensed Bold	26	75	27
Luckiesh & Moss (1940)	Memphis Light	11	86	49
	Memphis Medium	19	85	47
	Memphis Bold	25	92	41
Shaw (1969)	Gill Sans Roman	20	84	27
	Gill Sans Bold	31	100	26
	Plantin Roman	19	87	40
	Plantin Bold	32	101	38
Smither & Braun (1994)	Helvetica	17	82	25
	Helvetica Bold	27	90	25
Mansfield et al. (1996)	Times Roman	19	76	35
	Courier Bold	21	92	46
Sheedy et al. (2005)	Arial	17	81	26
	Arial Bold	26	91	27
	Verdana	17	84	32
	Verdana Bold	32	100	29
	Times New Roman	18	77	34
	Times New Roman Bold	30	97	14
	Georgia	19	81	41
	Georgia Bold	33	103	39
	Franklin Gothic Book	15	79	27
	Franklin Gothic Medium	22	83	27
	Franklin Gothic Demi	29	87	20
	Franklin Gothic Heavy	39	98	16

Appendix 2: Continued.

Scientific Study	**Typeface Tested**	**Parameter Values (% x-height)**		
		Stroke Width	**Letter Width**	**Letter Spacing**
Beveratou (2016)	Freight Sans Book	13	79	31
	Arial Regular	17	81	26
	Tiresias LPfont	26	83	29
Xiong et al. (2018)	Helvetica	17	82	25
	Times Roman	19	77	35
	Courier	16	83	50
	Maxular Rx Bold	28	107	78
Beier & Oderkerk (2019a)	Ovink Regular	14	83	-
	Ovink Semi Bold	29	97	-
	Ovink Ultra Black	53	111	-
Beier & Oderkerk (2019b)	Gill Sans Light	11	81	28
	KBH Text Regular	16	93	41

Appendix 3. Scientific review supplementary data. This appendix includes data from three studies that do not report statistical analyses.

Appendix 3a: Supplementary data: Roethlein, 1912. This table presents typefaces tested within the experimental study, and the reading performance data. Legibility is recorded as the average distance in centimetres from which the letters of a typeface can be identified (averaged over the twenty-six lowercase characters). Larger numbers correspond to higher legibility, meaning that the lowercase characters of a typeface can be identified from further away. The data is reproduced from Tables 3 and 6 in Roethlein (1912).

Typeface Tested	Distance (cm)
Century Old Style Bold	255.1
Century Old Style	228.0
Cheltenham Old Style Bold	233.4
Cheltenham Old Style	206.4
Cheltenham Condensed Bold	205.9

Appendix 3b: Supplementary data: Luckiesh & Moss, 1940. This table presents typefaces tested within the experimental study, and the reading performance data. Testing legibility with the Luckiesh-Moss visibility meter, the relative visibilities of three weights of Memphis are presented. The data is reproduced from Table 2 in Luckiesh and Moss (1940).

Typeface Tested	Relative Visibility
Memphis Bold	116.0
Memphis Medium	114.0
Memphis Light	100.0

Appendix 3c: Supplementary data: Beveratou, 2016. This table presents typefaces tested within the experimental study, and the reading performance data. Typefaces are scaled to have an x-height equivalent to Arial 16 point (underlined), therefore point sizes are included in the table. Words Read represents the total number of words read by the twenty-one partially sighted participants. Typefaces are presented in descending order according to Words Read data. Asterisks indicate typefaces analysed in Chapter 4. The data is reproduced from Table 1 in Beveratou (2016).

Typeface Tested	**Point Size**	**Words Read**
Tiresias LPfont*	15.5	491
Freight Text Book	18	458
Miniscule 6	16	447
Optima	18	443
Arial Regular*	16	440
Palatino	18	438
Times New Roman	18	429
Freight Micro Book	18	423
Century Gothic	15.5	421
Freight Sans*	18	406

Appendix 4. Twenty typefaces investigated and points accrued. The typefaces investigated in Chapter 6 are presented in alphabetical order alongside points accrued in 2010 and 2020. See Chapter 6 (section 6.2) for further information.

Typeface	**Typeface Distributors**		**Typography Manuals**	**History Books**	**Top 100 Lists**	**Total Points**
	2020	**2010**	**2010**	**2010**	**2010**	**2010**
Arial	2	2	1	1	0	4
Avenir	4	5	0	0	1	6
DIN	4	3	2	0	1	6
Franklin Gothic	2	1	4	4	1	10
Frutiger	4	4	6	1	1	12
Gill Sans	3	2	6	6	1	15
Helvetica	4	5	6	3	2	16
Neue Helvetica	4	4	1	2	0	7
Meta	3	2	3	1	2	8
Myriad	1	4	1	1	1	7
News Gothic	1	1	2	1	1	5
Officina	2	0	2	1	1	4
Optima	3	3	5	4	2	14
Rotis	2	3	0	0	2	5
Scala	1	1	6	1	1	9
Stone	1	0	2	2	2	6
Syntax	1	0	4	0	2	6
Thesis	3	1	0	1	2	4
Trade Gothic	4	4	0	0	1	5
Univers	4	4	6	6	2	18

LIST OF REFERENCES

Adobe Fonts, 2022. [online] Available at: <https://fonts.adobe.com/> [Accessed 15 April 2022].

Ambrose, G. and Harris, P., 2005. *Basic Design: Typography.* Lausanne: AVA Publishing SA.

Arditi, A., 1996. Typography, Print Legibility and Low Vision. In: B. Rosenthal and R. Cole, eds. *Remediation and Management of Low Vision*. St. Louis: Mosby.

Arditi, A., 2004. Adjustable Typography: An Approach to Enhancing Low Vision Text Accessibility. *Ergonomics*, 47(5), pp.469-482.

Arditi, A., 2018. *Making Text Legible: Designing for People With Partial Sight*. [pdf] New York: Lighthouse Guild. Available at: <https://lighthouseguild.org/effective-color-contrast-and-making-text-legible/> [Accessed 3 December 2021].

Arditi, A., Cagenello, R. and Jacobs, B., 1995a. Letter Strokewidth, Spacing, and Legibility. *Vision Science and its Applications*. Washington D.C.: Optical Society of America.

Arditi, A., Cagenello, R. and Jacobs, B., 1995b. Effects of Aspect Ratio and Spacing on Legibility of Small Letters. *Investigative Ophthalmology*, 36(4), pp.671.

Arditi, A. and Cho, J., 2005. Serifs and Font Legibility. *Vision Research*, 45(23), pp.2926-2933.

BIO.23 Biennial of Design 2012 [exhibition]. MAO I Museum of Architecture and Design, Ljubljana. 27 September-11 November 2012.

Baines, P. and Haslam, A., 2005. *Type and Typography.* 2nd ed. New York: Waston-Guptill Publications.

Barnes, C. and Melles, G., 2007. Managing Interdisciplinarity: A Discussion of the Contextual Review in Design Research. *Proceedings of the International Association of Societies of Design Research.* Hong Kong: Hong Kong Polytechnic University.

Beier, S., 2012. *Reading Letters: Designing for Legibility*. Amsterdam: BIS Publishers.

Beier, S., 2016. Letterform Research: An Academic Orphan. *Visible Language*, 50(2), pp.64-79.

Beier, S., 2018. The Relationship Between Stroke Weight and Letter Width. Association Typographique Internationale, *ATypI Antwerp 2018.* Antwerpen, Belgium, 11-15 September 2018.

Beier, S. and Dyson, M., 2014. The Influence of Serifs on ‘h’ and ‘i’: Useful Knowledge from Design-Led Scientific Research. *Visible Language*, 47(3), pp.74-95.

Beier, S. and Oderkerk, C.A.T., 2019a. Smaller Visual Angles Show Greater Benefit of Letter Boldness Than Larger Visual Angles. *Acta Psychologica,* 199, pp.1-8.

Beier, S. and Oderkerk, C.A.T., 2019b. The Effect of Age and Font on Reading Ability. *Visible Language,* 53(3), pp.50-68.

Bennett, A., 2006. *Design Studies: Theory and Research in Graphic Design.* New York: Princeton Architectural Press.

Berger, C. 1948. Some Experiments on the Width of Symbols as Determinant of Legibility. *Acta Ophthalmologica*, 26(4), pp.517-550.

Bernard, J-B., Aguilar, C. and 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(4), pp.1-25.

Bernard, J-B., Kumar, G., Junge, J. and Chung, S., 2013. The Effect of Letter-Stroke Boldness on Reading Speed in Central and Peripheral Vision. *Vision Research*, 84, pp.33-42.

Bernau, K., 2005. *Neutral*. The Hague: Macdonald/SSN and Nijmegen.

Bessemans, A., 2012. *Letterontwerp voor kinderen met een visuele functiebeperking*. PhD. Leiden University and Hasselt University.

Bessemans, A., 2016a. Matilda: A Typeface for Children With Low Vision. In: M.C. Dyson and C.Y. Suen, eds. *Digital Fonts and Reading*. Singapore: World Scientific.

Bessemans, A., 2016b. Typeface’s for Children’s Reading. *TIJDSCHRIFT VOOR MEDIAGESCHIEDENIS,* 19(2), pp.1-9.

Beveratou, E., 2016. The Effect of Type Design and Typesetting on Visually Impaired Readers. In: M.C. Dyson and C.Y. Suen, eds. *Digital Fonts and Reading*. Singapore: World Scientific.

Bigelow, C., 2019. Typeface Features and Legibility Research. *Vision Research*, 165, pp.162-172.

Bigelow & Holmes, 2015. *On Font Weight.* [online] Available at: <https://bigelowandholmes.typepad.com/bigelow-holmes/2015/07/on-font-weight.html> [Accessed 17 December 2020].

Bouma, H., 1970. Interaction Effects in Parafoveal Letter Recognition. *Nature*, 226, pp.177-178.

Bringhurst, R., 2002. *The Elements of Typographic Style.* 2nd ed. Vancouver: Hartley & Marks Publishers.

Burmistrov, I., Zlokazova, T., Ishmuratova, I. and Semenova, M., 2016. Legibility of Light and Ultra-Light Fonts: Eyetracking Study. In: NordiCHI '16, *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. Gothenburg, Sweden, October 2016. New York: Association for Computing Machinery.

CNIB, 2006. *Clear Print Accessibility Guidelines*. Toronto: CNIB.

CNIB, 2020. *Clear Print Accessibility Guidelines.* [pdf] Toronto: Canadian National Institute for the Blind (CNIB). Available at: <https://www.cnib.ca/sites/default/files/2020-08/Clear%20Print%20Guidelines%202020.pdf> [Accessed 24 April 2021].

Cabarga, L., 2006. *Learn FontLab Fast*. Los Angeles: Iconoclassics.

Campbell, K.A., Cutler, F., McDonald, R., Putt, C., Rewak, M., Strong, G. and Whitton, H., 2005. *CNIB/OCAD Typographic Legibility Research Project: Clear Print Report*. Toronto: CNIB/OCAD Research.

Candy, L., 2006. Practice Based Research: A Guide. *Creativity and Cognition Studios Report*, 1, pp.1-19.

Candy, L. and Edmonds, E., 2018. Practice-Based Research in the Creative Arts: Foundations and Futures from the Front Line. *LEONARDO,* 51(1), pp.63-69.

Carter, R., Day, B. and Meggs, P., 2007. *Typographic Design: Form and Communication*. 4th ed. New Jersey: John Wiley and Sons.

Carvalho, L. and Dong, A., 2009. Legitimating Design: A Sociology of Knowledge Account of the Field. *Design Studies,* 30(5), pp.483-502.

Cheng, K., 2005. *Designing Type.* New Haven: Yale University Press.

Chung, S., 2012. Dependence of Reading Speed on Letter Spacing in Central Vision Loss. *Optometry and Vision Science*, 89(9), pp.1288-1298.

Clarkson, J., Coleman, R., Keates, S. and Lebbo, C., ed. 2003. *Inclusive Design: Design for the Whole Population.* London: Springer.

Craig, J., Scala, I.K. and Bevington, W., 2006. *Designing With Type: The Essential Guide to Typography.* 5th ed. New York: Watson-Guptill Publications.

Cramsie, P., 2010. *The Story of Graphic Design.* London: The British Library.

Cross, N., 2001. Designerly Ways of Knowing: Design Discipline versus Design Science. *Design Issues,* 17(3), pp.49-55.

Cross, N., 2007. *Designerly Ways of Knowing.* Basel: Birkhäuser.

Delve Fonts, 2021. *Maxular*. [online] Available at: <https://delvefonts.com/fonts/maxular#buy> [Accessed 24 February 2021].

Devroye, L., 2022. *Paul Shaw*. [online] Available at: <http://luc.devroye.org/fonts-33597.html#:~:text=Paul%20Shaw&text=Paul%20Shaw's%20choice%20of%20100,Helvetica%2C%20Aldus%20Manutius'%20roman.> [Accessed 15 April 2022].

Dodd, R., 2006. *From Gutenberg to OpenType: An Illustrated History of Type From the Earliest Letterforms to the Latest Digital Fonts.* Vancouver: Hartley & Marks Publishers.

Dyson, M., 1999. Typography Through the Eyes of a Psychologist. *Hyøen*, 1, pp.5-13.

Dyson, M.C., 2013. Where Theory Meets Practice: A Critical Comparison of Research into Identifying Letters and Craft Knowledge of Type Design. *The Design Journal*, 16(3), pp.271-294.

Dyson, M. and Beier, S., 2016. Investigating Typographic Differentiation. *Information Design Journal*, 22(1), pp.3-18.

Ellison, A., 2006. *The Complete Guide to Digital Type: Creative Use of Typography in the Digital Arts.* London: Laurence King Publishing Ltd.

Eskilson, S.J., 2007. *Graphic Design: A New History.* London: Laurence King.

Evamy, M. and Roberts, L., 2004. *In Sight: A Guide to Design With Low Vision in Mind.* Mies: RotoVision.

Felici, J., 2012. *The Complete Manual of Typography: A Guide to Setting Perfect Type.* 2nd ed. Berkeley: Peachpit.

FontLab, 2006. *FontLab Studio: User's Manual for Macintosh*. Port Angeles: FontLab.

Fonts.com, 2020a. *Best Sellers*. [online] Available at: <https://www.fonts.com/browse/best-sellers> [Accessed 6 August 2020].

Fonts.com, 2020b. *Eurostile LT Typeface Story*. [online] Available at: <https://www.fonts.com/font/linotype/eurostile-lt/story> [Accessed 7 August 2020].

Fonts.com, 2020c. *Futura Typeface Story*. [online] Available at: <https://www.fonts.com/font/linotype/futura/story> [Accessed 7 August 2020].

Fonts.com, 2020d. *ITC Avant Garde Gothic Typeface Story*. [online] Available at: <https://www.fonts.com/font/itc/itc-avant-garde-gothic/story> [Accessed 7 August 2020].

Fonts.com, 2020e. *Kabel Typeface Story*. [online] Available at: <https://www.fonts.com/font/linotype/kabel/story> [Accessed 7 August 2020].

FontShop, 2020. *Best Selling Fonts*. [online] Available at: <https://www.fontshop.com/best-selling-fonts> [Accessed 6 August 2020].

FontShop, 2022. *100 Best Typefaces of All Time.* [online] Available at: <https://www.fontshop.com/people/fontshop-team/fontlists/100-best-typefaces-of-all-time> [Accessed 15 April 2022].

Frayling, C., 1993. Research in Art and Design. *Royal College of Art Research Papers,* 1(1).

Frensch, N., 2003. *Read Regular™.* London: Natasha Frensch.

Friedman, K., 2008. Research Into, By and For Design. *Journal of Visual Arts Practice,* 7(2), pp.153-160.

Garvey, P., Pietrucha, M., and Meeker, D., 1998. Clearer Road Signs Ahead. *Ergonomics in Design*, 6, pp.7-11.

Godin, D. and Zahedi, M., 2014. Aspects of Research Through Design: A Literature Review. In: Y. Lim, K. Niedderer, J. Redström, E. Stolterman and A. Valtonen, eds. *Design's Big Debates - DRS International Conference*, Umeå, Sweden, 16-19 June 2014.

Gray, C. and Malins, J., 2004. *Visualizing Research: A Guide to the Research Process in Art and Design*. Surrey, England: Ashgate Publishing Limited.

HSBC UK, 2022. *Visual Impairment*. [online] Available at: <https://www.hsbc.co.uk/accessibility/visually-impaired/#braille> [Accessed 24 April 2022].

Haley, A., 2020. *Uses of Bold Type in Text.* [online] Available at: <https://www.fonts.com/content/learning/fontology/level-2/text-typography/bold-type-in-text#:~:text=Bold%20type%20grew%20out%20of,used%20to%20set%20text%20copy> [Accessed 17 December 2020].

Hand, D.J., 2008. *Statistics: A Very Short Introduction*. Oxford: Oxford University Press.

Hedlich, C., Barstow, E. and Vogtle, L.K., 2018. Age-Related Macular Degeneration and Reading Performance: Does Font Style Make a Difference? *Journal of Visual Impairment & Blindness,* 112(4), pp.398-402.

Heylighen, A., Cavallin, H. and Bianchin, M., 2009. Design in Mind*. Design Issues*, 25(1), pp.94-105.

Hill, W., 2010. *The Complete Typographer: A Foundation Course for Graphic Designers Working With Type.* 3rd ed. London: Thames & Hudson Ltd.

Hollis, R., 2001. *Graphic Design: A Concise History.* London: Thames & Hudson.

Jones, M. and Mewhort, D., 2004. Case-Sensitive Letter and Bigram Frequency Counts From Large-Scale English Corpora. *Behavior Research Methods, Instruments, & Computers*, 36, pp.388–396.

Jury, D., 2004. *About Face: Reviving the Rules of Typography.* Mies: RotoVision.

Kitchel, J.E., 2004. *Large Print: Guidelines for Optimal Readability and APHont a Font for Low Vision*. Louisville: American Printing House for the Blind.

Klanten, R., Bourquin, N., Tissot, T. and Ehmann, S. eds., 2008. *Data Flow: Visualizing Information in Graphic Design*, Berlin: Gestalten.

Knuth, D.D., 1986. *The METAFont Book.* Boston: American Mathematical Society, Addison-Wesley.

Legge, G.E., 2007. *Psychophysics of Reading in Normal and Low Vision.* London: Lawrence Erlbaum Associates.

Legge, G.E., 2016. Reading Digital With Low Vision. *Visible Language*, 50(2), pp.1-13.

Legge, G.E. and Bigelow, C., 2011. Does Print Size Matter for Reading? A Review of Findings from Vision Science and Typography. *Journal of Vision*, 11(5), pp.1-22.

Linotype, 2020. *Best Selling Fonts*. [online] Available at: <https://www.linotype.com/catalog/bestsellers.html> [Accessed 6 August 2020].

Linotype, 2021. *Neue Helvetica*. [online] Available at: <https://www.linotype.com/1266/neue-helvetica-family.html> [Accessed 8 April 2021].

Lovie-Kitchen, J. and Whittaker, A., 1998. Relative-Size Magnification Versus Relative-Distance Magnification: Effect on the Reading Performance of Adults With Normal and Low Vision. *Journal of Visual Impairment and Blindness,* 92(7), pp.433-446.

Luckiesh, M. and Moss, F.K. 1939. The Visibility and Readability of Printed Matter. *Journal of Applied Psychology,* 23, pp.645-659.

Luckiesh, M. and Moss, F.K., 1940. Boldness as a Factor in Type-Design and Typography. *Journal of Applied Psychology,* 24(2), pp.170-183.

Lund, O., 1999. *Knowledge Construction in Typography: The Case of Legibility Research and the Legibility of Sans Serif Typefaces.* PhD. The University of Reading.

Lupton, E., 2004. *Thinking With Type: A Critical Guide for Designers, Writers, Editors, and Students.* New York: Princeton Architectural Press.

Mansfield, J.S., Ahn, S.J., Legge, G.E. and Luebker, A., 1993. A New Reading Acuity Chart for Normal and Low Vision. *Ophthalmic and Visual Optics/Noninvasive Assessment of the Visual System Technical Digest*, 3, pp.232-235.

Mansfield, S., Legge, G. and Bane, M., 1996. Psychophysics of Reading XV: Font Effects in Normal and Low Vision. *Investigative Ophthalmology and Visual Science*, 37(8), pp.1492-1501.

Mariotti, S.P., 2012. *Global Data on Visual Impairments 2010*. Geneva: World Health Organization.

McLean, R., 1980. *The Thames and Hudson Manual of Typography.* London: Thames & Hudson.

Meggs, P.B., 1998. *A History of Graphic Design.* 3rd ed. New York: John Wiley & Sons, Inc.

Meggs, P.B. and Purvis, A.W., 2006. *Meggs' History of Graphic Design*. Hoboken New Jersey: John Wiley & Sons, Inc.

Morris, R., Aquilante, K., Yager, D. and Bigelow, C., 2002. Serifs Slow RSVP Reading at Very Small Sizes, but Don't Matter at Larger Sizes. *Society for Information Display International Symposium Digest of Technical Papers*, 33, pp.1-4.

MyFonts, 2020. *Best Sellers*. [online] Available at: <https://www.myfonts.com/bestsellers/> [Accessed 6 August 2020].

MyFonts, 2022. *Tiresias*. [online] Available at: <https://www.myfonts.com/fonts/bitstream/tiresias> [Accessed 25 April 2022].

Norton, T., Corliss, D. and Bailey, J., 2002. *The Psychophysical Measurement of Visual Function*. Woburn: Elsevier Science.

Polanyi, M., 1976. *The Tacit Dimension.* Garden City: Doubleday.

Poynor, R., 2003. *No More Rules: Graphic Design and Postmodernism.* New Haven: Yale University Press.

Purvis, A.W. and Le Coultre, M.F., 2003. *Graphic Design 20th Century.* Amsterdam: BIS Publishers.

RNIB, 2017. *Top Tips for Creating Accessible Print Documents*. [pdf] RNIB. Available at: <https://www.rnib.org.uk/sites/default/files/Top_Tips_Creating_accessible_print_documents.pdf [Accessed 24 April 2021].

RNIB, 2018. *Eye Health and Sight Loss Stats and Facts*. [pdf] RNIB. Available at: <https://www.rnib.org.uk/sites/default/files/Eye%20health%20and%20sight%20loss%20stats%20and%20facts.pdf> [Accessed 18 June 2020].

RNIB, 2019. *Key Information and Statistics on Sight Loss in the UK*. [online] Available at: https://www.rnib.org.uk/professionals/knowledge-and-research-hub/key-information-and-statistics> [Accessed 18 June 2020].

RNIB, 2020. *Equality Act Compliance*. [online] Available at: <https://www.rnib.org.uk/services-we-offer-advice-professionals/equality-act-compliance> [Accessed 18 June 2020].

RNIB and ISTD, 2007. *Inclusive Design: Clear and Large Print Best Practice Guide for Designers.* Taunton: International Society of Typographic Designers.

RTD2017, 2017. *What is RTD?* [online] Available at: <https://www.researchthroughdesign.org/2017/whatisrtd/> [Accessed 8 April 2021].

Radner, W., 2017. Reading Charts in Ophthalmology. *Graefe's Archive for Clinical and Experimental Ophthalmology,* 255(8), pp.1465-1482.

Roethlein, B.E., 1912. The Relative Legibility of Different Faces of Printing Types. *American Journal of Psychology,* 23, pp.1-36.

Rubin, G.S., Feely, M., Perera, S., Ekstrom, K. and Williamson, E., 2006. The Effect of Font and Line Width on Reading Speed in People With Mild to Moderate Vision Loss. *Ophthalmic & Physiological Optics: The Journal of The British College of Ophthalmic Opticians (Optometrists),* 26(6), pp.545-554.

Rubin, G.S. and Turano, K., 1992. Reading Without Saccadic Eye Movement. *Vision Research,* 32(5), pp.895-902.

Russell-Minda, E., Jutai, J., Strong, J., Campbell, K., Gold, D., Pretty, L. and Wilmot, L., 2007. The Legibility of Typefaces for Readers With Low Vision: A Research Review. *Journal of Visual Impairment and Blindness,* 101(7),402-415.

Shaw, A., 1969. *Print for Partial Sight: A Report to the Library Association Sub-Committee on Books for Readers With Defective Sight.* London: Library Association.

Shaw, P., 2010. *Ten Typefaces of the Decade.* [online] Available at: <https://www.printmag.com/paul-shaw/ten-typefaces-of-the-decade/> [Accessed 15 April 2022].

Sheedy, J., Subbaram, M., Zimmerman, A. and Hayes, J., 2005. Text Legibility and the Letter Superiority Effect. *Human Factors*, 47(4), pp.797-815.

Smither, J.A. and Braun, C.C., 1994. Readability of Prescription Drug Labels by Older and Younger Adults. *Journal of Clinical Psychology in Medical Settings,* 1, pp.149-159.

Storkerson, P., 2008. Is Disciplinary Research Possible in Communication Design? *Design Research Quarterly,* 3(2), pp.1-8.

Strickler, Z., 1999. Elicitation Methods in Experimental Design Research. *Design Issues,* 15(2), pp.27-39.

Subbaram, V.M., 2004. *Effect of Display and Text Parameters on Reading Performance*, PhD. Ohio State University.

Tarita-Nistor, L., Lam, D., Brent, M., Steinbach, M. and Gonzalez, E., 2013. Courier: A Better Font for Reading With Age-Related Macular Degeneration. *Canadian Journal of Ophthalmology*, 48(1), pp.56-62.

Tinker, M.A., 1963. *Legibility of Print*. Ames: Iowa State University Press.

Tinker, M.A., 1965. *Bases for Effective Reading*. Minneapolis: University of Minnesota Press.

Tracy, W., 1986. *Letters of Credit: A View of Type Design.* Boston: David R. Godine.

Twyman, M., 1993. The Bold Idea: The Use of Bold-looking Types in the Nineteenth Century. *The Journal of the Printing Historical Society*, 22, pp.107-143.

UKAAF, 2019. *Creating Clear Print and Large Print Documents: Guidance from UKAAF*. [pdf] UK Association for Accessible Formats (UKAAF). Available at: <https://www.ukaaf.org/wp-content/uploads/2020/03/G003-UKAAF-Creating-clear-print-and-large-print-documents-v2.pdf> [Accessed 15 April 2022].

Unger, G., 2005. *While You're Reading*. New York: Mark Batty.

United Nations, 2020. *World Population Ageing 2019*. [pdf] United Nations. Available at: <https://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2019-Report.pdf> [Accessed 15 April 2022].

Von Ompteda, K., 2008. The Role of Typeface Design Within the Scientific Study of Legibility. Association Typographique International, *ATypI 08 St Petersburg (TypeTech by ATypI).* St. Petersburg, Russia, 17-21 September 2008.

Von Ompteda, K., 2009a. Innovation in Inclusive Typography: A Role for Design Research. In: Helen Hamlyn Centre, *Include 2009: International Conference on Inclusive Design*. London, United Kingdom, 5-8 April 2009.

Von Ompteda, K., 2009b. Inclusive Typography: Toward the Integration of Vision Science and Communication Design. University of Cambridge, Law Faculty, *ICG09 – The Interdisciplinary Graduate Conference 2009*. Cambridge, United Kingdom, 25-26 June 2009.

Von Ompteda, K., 2011a. Typo Transparencies. *Grafik Magazine,* 192, pp.83-97.

Von Ompteda, K., 2011b. Visualising Type. *TYPO London*. London, United Kingdom, 20-22 October 2011.

Von Ompteda, K., 2019a. Data Manifestation. In: T. Triggs and L. Atzmon, eds. *The Graphic Design Reader*. London: Bloomsbury Press.

von Ompteda, K., 2019b. Data Manifestation: Merging the Human World and Global Climate Change. In: *IEEE VIS 2019 Arts Program*. Vancouver, Canada, 21-25 October 2019.

Walker, S., 2017. Research in Graphic Design. *The Design Journal*, 20(5), pp.549-559.

Waller, R., 2011. *Technical Paper 10. The Clear Print Standard: Arguments for a Flexible Approach.* [pdf] Simplification Centre. Available at: <https://www.simplificationcentre.org.uk/reports2/technical-paper-10-the-clear-print-standard-arguments-for-a-flexible-approach> [Accessed 17 December 2020].

Xiong, Y., Lorsung, E.A., Mansfield, J.S., Bigelow, C. and Legge, G.E., 2018. Fonts Designed for Macular Degeneration: Impact on Reading. *IOVS,* 59(10), pp.4182-4189.

Yager, D., Aquilante, K. and Plass, R., 1998. High and Low Luminance Letters, Acuity Reserve, and Font Effects on Reading Speed. *Vision Research,* 38(17), pp.2527-2531.

Yau, N., 2013. *Data Points: Visualization That Means Something*, Indianapolis: John Wiley & Sons, Inc.

Zar, J.H. 1999. *Biostatistical Analysis*. 4th ed. New Jersey: Prentice-Hall, Inc.