



Typeface features and legibility research

Charles Bigelow

Cary Graphic Arts Collection, Rochester Institute of Technology, 90 Lomb Memorial Drive, Rochester, NY 14623-5604, USA



ABSTRACT

In the early 20th century, reading researchers expressed optimism that scientific study of reading would improve the legibility of typefaces. Font-making was, however, complex, expensive and impractical for reading research, which was therefore restricted to standard commercial fonts. The adoption of computer typography in legibility studies makes the measurement, modification and creation of experimental fonts easier, while display of text on computer screens facilitates reading studies. These technical advances have spurred innovative research. Some studies continue to test fonts for efficient reading in low vision as well as normal vision, while others use novel fonts to investigate visual mechanisms in reading. Some experimental fonts incorporate color and animation features that were impractical or impossible in traditional typography. While it is not clear that such innovations will achieve the optimistic goals of a century ago, they extend the investigation and understanding of the nature of reading.

1. Introduction

The optimistic goal of studying and improving letter forms is expressed in the words of early reading researchers,

“For more than thirty centuries, the characters used by mankind to record its thoughts have evolved almost without method, by force of circumstances. The result is that our modern writing, from that of the young schoolchild to the most elegant typography, constitutes an offense to good sense and is tolerated only because of the habit of centuries that transmits it from generation to generation.” (Javal, 1905).

“Certainly the letter-forms that have come down to us through the ages have never been pruned to meet the reader’s needs.... There is not the slightest doubt that forms can be devised which will be much more legible than these ancient traditional symbols.” (Huey, 1908). “Every reader has observed that all of these variants of letter-forms are not equally legible – an observation which raises the theoretical question: What are the factors upon which legibility depends? And the practical question: How should one proceed if one set out to improve the legibility of printed letters?” (Roethlein, 1912).

Yet, early on, Richard L. Pyke expressed skepticism:

“The hypothesis is here put forward that extremely large typographic differences must be present before it is possible to say that there is any difference in the objective legibility of types.... Four times as many writers have measured legibility as have defined it. Three out of every four writers have been attempting to measure something the exact nature of which they have not paused to examine.” (Pyke, 1926).

Later, surveying four decades of legibility research, Miles Tinker observed that progress had been slow.

“Before the nineteenth century, the main concern was with esthetic appearance of print. With improved technology of printing, two additional factors entered the picture: Economy of printing and traditional practices. Because of these practices and views, a truly scientific typography has been slow in developing.” (Tinker, 1963, p.3).

By the early 21st century, however, reading research had rebounded in general and in particular by application of psychophysical theory and methodology.

“Legibility is a good example of a psychophysical variable, such as color or brightness, which is dependent on physical stimulus properties, but is fundamentally determined by characteristics of visual processing.” (Legge, 2006, p. 108).

2. The making of type, traditional and modern

To study legibility, it is helpful to know something about how type was and is made, and its terminology. In the 500 year era of metal typography, a “font” was a set of cast metal letters and characters of a given size and style. A “typeface” comprised the shaped surfaces of the metal that were inked and impressed on paper in printing, and hence the look or design of a specific style. Today, “font” means a typeface or a font – the look or the digital file.

E-mail address: chuck.bigelow@gmail.com.

<https://doi.org/10.1016/j.visres.2019.05.003>

Received 30 June 2018; Received in revised form 29 April 2019; Accepted 2 May 2019

Available online 16 November 2019

0042-6989/ © 2019 Elsevier Ltd. All rights reserved.



Fig. 1. Punches of a Baskerville font, cut by John Handy circa 1757.

2.1. Traditional type

From the 15th century to the end of the 19th century, type was made by skilled artisans who hand-cut letter images in relief on steel “punches” for each letter of each size of each style of type. See Fig. 1. Finished punches were hammered into copper blanks to make recessed impressions, called matrices, from which type was hand-cast (Dreyfus, 1995). The art of hand punch-cutting is still practiced by a few highly skilled letter engravers, who are sculptors of miniature forms (Gable & Paput, 2016).

Punch-cutters worked at the actual size of a font and thus saw, tested, and corrected letter forms at the same size as readers saw them. Since the end of the 19th century, however, type designs have been created at larger sizes, around 75–150 mm, to be scaled mechanically, photographically, or by computer to small reading sizes.

Giambattista Bodoni (1818) stated that a typeface is more beautiful when it has regularity (including formal repeatability and symmetry), clarity (including neatness and smoothness), good taste (including simplicity and pleasing chiaroscuro of black and white); and grace, a quality Bodoni himself found difficult to describe precisely. Most text typefaces for continuous reading manifest those aesthetic principles as: even spacing of strokes within letters and between letters, resulting in a visually regular rhythm; regular thicknesses of strokes (both for thick and thin strokes); replication of similar forms such as serifs or bowls of ‘b’ and ‘p’; consistent gray tone so no letter looks darker or lighter than others; visually regular horizontal alignments of letters along the baseline, x-height, and capital height.

2.2. Digital type

Most software fonts define letter shapes as mathematical outlines made of straight lines and curves called “splines.” The outline letter shapes in computer fonts can be linearly scaled to any practical size. In computer display or printing, the outlines after scaling are rasterized – converted into pixels (“picture elements”) – for digital output. An example of an outline and rasterized letter is in Fig. 2.

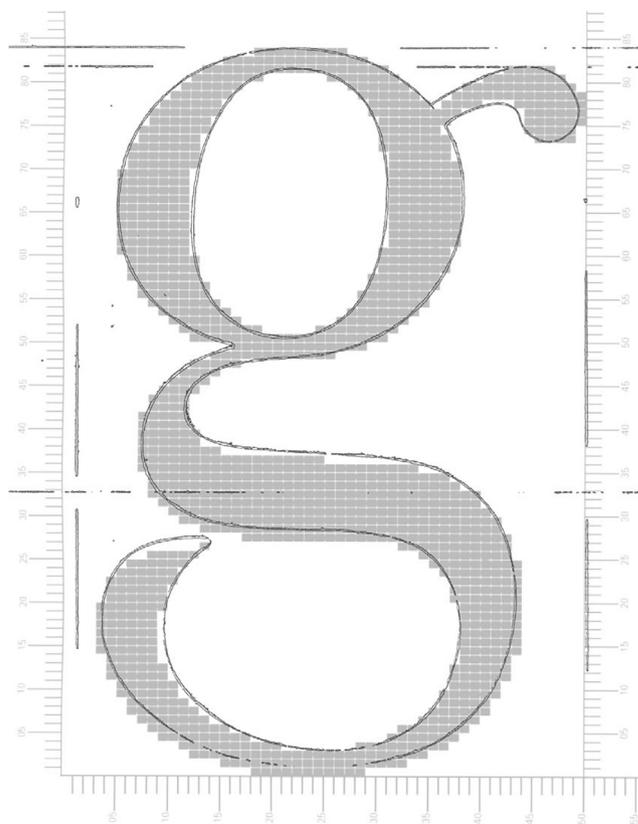


Fig. 2. Digital Baskerville g. Spline outlines with rasterized bitmap.

Compared to metal and photo fonts of earlier eras, software fonts are more useful in reading research because they are easier and cheaper to create or modify; easier to measure precisely in sizes and features; easier to test on screens or in print; easier to reproduce; more portable across digital media and computer systems.

The majority of type designers now use software tools to create fonts by drawing large letter outlines on computer screens (Fig. 3). Font software tools include: FontForge, FontLab, Fontographer, Glyphs, OTMaster, Robofont, TypeTool (Tools, 2019) and Metafont (Knuth, 1986). Cheng (2006) explains type design basics, including classifications, feature variables, and design principles, with tutorial exercises for creating fonts.

3. Topics in legibility research

The main goal of 20th century legibility research was to determine which features and typographical variations affected legibility in publishing, signage, and labeling. Wendt (1994), Lund (1999), Beier (2012), and Bigelow (2016) include histories, and discussions of legibility research on salient features and parameters of size, width, weight, italic, serifs, contrast, and capitals.

4. Size

The effect of print size on reading, particularly the minimum size for easy reading, is perhaps the oldest forms of legibility research because print size is important economically as well as perceptually. Printers prefer smaller sizes that occupy less area for a given amount of text and thus use less paper and reduce printing costs. Readers usually prefer print sizes large enough for easy reading. Hence, there are compromises between opposing economic and perceptual preferences. (Here, the terms “print size” and “type size” are used equivalently, although “type size” may refer to screen size as well as to print.)

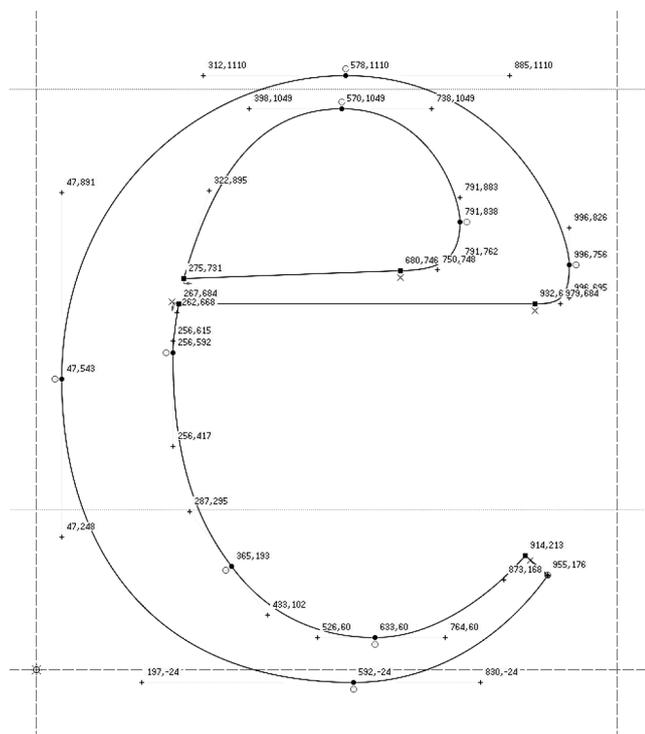


Fig. 3. Baskerville letter e in Bézier spline curves. Solid lines define the outline of the letter. Dots with coordinate pairs mark important points or “knots” on curves, tangents, and corners. Dashed vertical lines mark “sidebearing” boundaries of letter width. The dashed horizontal line near bottom is the font baseline; the dotted horizontal line near top is the font x-height.

Legge and Bigelow (2011) offer reasons for using x-height as a standard measure of character size in vision research. The x-height of a font is measured from the implicit “baseline” on which letters stand, to the top of the lowercase ‘x’, as illustrated in Fig. 4.

Some researchers measure the heights of “small” letters like ‘a’, ‘e’, and ‘o’, which are very slightly taller than ‘x’, but at common type sizes for reading continuous text, say, 9 point to 16 point, the difference is negligible for most purposes.

The x-height of a font can be expressed as a fraction of the full font body size and can be determined precisely for digital fonts. For instance, in an OpenType font in which the coordinates of the body size, also called typographic “em,” are on a 1000 × 1000 units Cartesian grid, the x-height of Times Roman is 450 units and the x-height fraction is 0.45.

When reading distance and physical x-height are known, the visual angle of x-height subtended at the eye can be calculated as a single psychophysical measure.

Legge (2006) summarizes findings of several experiments on character size on reading speed, finding that speed is roughly constant in a ten-fold range between 0.2 and 2.0 degrees of visual angle of x-height. At sizes below 0.2 degrees of visual angle, reading speed falls quickly, while above 2.0 degrees, reading speed slows gradually. Legge defines “critical print size (CPS)” as the “smallest print size below which reading speed begins to decline sharply,” and notes that, “Across studies, a consensus value for the CPS for normally sighted readers is 0.2 degrees (12 min-arc).”



Fig. 4. Comparison of x-heights. From left: big x-height (Lucida Bright); medium x-height (Times Roman); small x-height (Centaur).

For example, the x-height of 9 point Times Roman at a reading distance of 40 cm subtends approximately 0.20 degrees of visual angle. The x-height of 90 point Times at 40 cm distance subtends approximately 2.0 degrees of visual angle. This span is roughly the type size range of a print newspaper, from running text to large headlines.

Physical type size on a computer, tablet or smart phone screen may differ from menu font point size depending on screen resolution, or setting if display is adjustable, and on word processing application. Hence, measurement of actual, on-screen physical type size is often more accurate than the font point size in a word processing menu.

4.1. Minimum size recommendations

Over a century, legibility studies have suggested that the smallest print sizes that can be read easily have a lowercase x-height around 1.5 mm. At a reading distance of 40 cm, 1.5 mm subtends 0.21 degrees of visual angle.

Griffing and Franz (1896) concluded that size is the most important factor related to visual fatigue and recommended that type sizes below 1.5 mm in x-height not be used.

Javal (1905) stated that, “In France, it is nine [point] that is most used for books and main articles in newspapers.” The 9 point type in Javal’s book has an x-height of 1.5 mm.

Huey (1908) summarized the findings of Griffing and Franz (1896) and an 1881 study by Weber, and recommended x-heights between 1.5 mm and 2 mm for fast reading and quick recognition.

Roethlein (1912) tested legibility of sixteen text typefaces at 10 point. Their exact x-heights are difficult to determine today because most of the studied typefaces are no longer in use, but those that Roethlein found most legible had x-heights ranging roughly from 1.7 mm to 1.9 mm, while the least legible had x-heights around 1.5 mm.

The British Association for the Advancement of Science (1913) recommended minimum type x-height of 1.58 mm for schoolchildren over 12 years old.

Ovink (1938) stated that “All authors up till now are holding the view that the optimal size of printing types is about 9 or 10 point, i.e. at least 1.5 mm x-height, preferably a little more.”

Tinker (1963) summarized several studies by Paterson & Tinker, which found that print sizes of 10 and 11 point were read most quickly. Insofar as the x-heights and reading distances can be estimated today, the average was approximately 1.5 mm.

Legge and Bigelow (2011) measured x-heights of running text type in newspapers and books, and found that angular x-height in these “ecological” contexts instead of in laboratory studies ranged from 0.20 to 0.28 degrees. The mean for newspapers was 0.23 degrees, and for hardback books, 0.24 degrees. That mean print sizes in publishing are somewhat larger than critical print size found in laboratory studies may indicate compensations for the range of visual acuities of average readers.

4.2. Linear scale

Nearly all scaling of computer software fonts is linear: letter forms are multiplied by a fixed constant in both dimensions, preserving geometric similarity. In reading studies, linear scaling provides simplicity, economy, and predictability: letter shapes and proportions are invariant through all sizes. As an example, in the MNREAD acuity chart used to test the effect of different print sizes on reading performance, the Times Roman type is linearly scaled for all sizes (Mansfield, Legge, Luebker, & Cunningham, 2006).

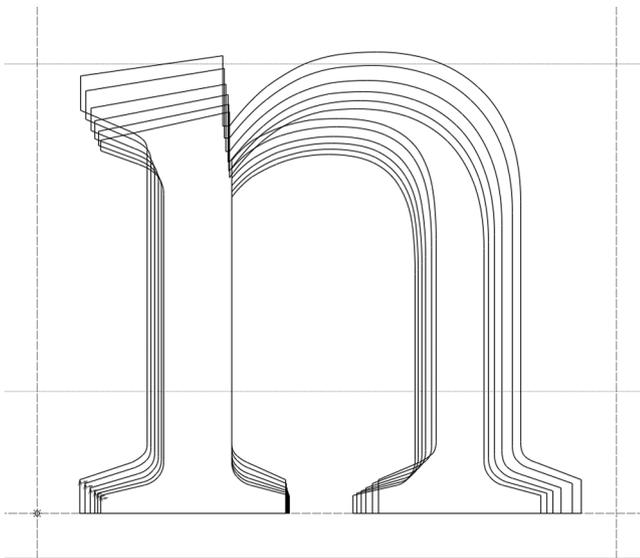


Fig. 5. Six optical scales of a letter 'n' in the "Sitka" typeface series. The tallest, widest, and thickest outline is for the smallest size-specific scale ("Sitka Small"); the shortest, narrowest, and thinnest outline is for the largest size-specific scale ("Sitka Banner").

4.3. Optical scale or size-specific design

Traditional hand punch-cutters as well as some 20th century type foundry and manufacturers of composing machines adjusted the proportions of letters depending on typographic scale. Types cut for large sizes were different from types cut for small sizes. This has been termed "optical scaling" (Carter, 1984/1937) or "size-specific design" (Larson & Carter, 2016).

As Carter put it, "The good punchcutters varied the design, or at any rate the functional features of it, to suit the scales on which they worked. They did so instinctively because they corrected their work by eye."

Fonts designed for small scale – that is, for small print sizes – look bigger and stronger because they have greater x-height fractions of body size, relatively wider letters, more space between letters, less difference between hairline and stem thicknesses, and thicker hairlines and serifs. Fonts designed for large scale look smaller and more delicate because they have smaller x-height fractions, are relatively narrower, have less space between letters, and thinner hairlines and serifs (see Fig. 5 and Fig. 15).

5. Type features

Typefaces are ensembles of graphical features and distinctions, which include: roman versus italic; normal versus light or bold weight; normal width versus condensed (narrow) or extended (wide) aspect ratios. Other variations include: presence or absence of serifs (seriffed versus sans-serif); large or small ratio of thick to thin strokes (high versus low "typographic contrast"); proportional-spacing (different letter widths) versus monospacing (typewriter-like fixed-widths); capitals versus lowercase. Several such distinctions can be seen in three familiar type families: Times Roman, Helvetica, and Courier, shown in Fig. 6.

Times and Courier are both seriffed with nearly identical x-heights, but Times has high contrast between thick and thin features, is proportionally spaced, and has a narrow lowercase. Courier has low contrast, is monospaced, and is wider than Times. Helvetica is sans-serif, has a larger x-height than either Times or Courier, has low contrast, is proportionally spaced, and is narrower than Courier but wider than Times. (The common sans-serif typeface "Arial" is very similar to Helvetica.)

Times: ABCEGHabcde fghijklmno

Helvetica: ABCEGHabcde fghijklmno

Courier: ABCEGHabcde fghijklmno

Fig. 6. Times Roman, Helvetica, Courier typefaces at same point size.

Differences in features may affect legibility, although it can be difficult to determine the features that are more influential, which may depend on reader vision as well as on type design. Mansfield, Legge, and Bane (1996) found that readers with normal vision read Times Roman 5% faster than Courier Bold at moderate print sizes (above critical print sizes for the respective fonts), but read Courier Bold up to twice as fast as Times at tiny sizes (below critical print sizes). Readers with low vision read Courier faster than Times at both moderate and tiny sizes. Hence, legibility may vary according to type size, width, and boldness, and be strongly affected by reader vision, indicating that legibility is not solely a graphical characteristic. Tarita-Nistor et al. (2013) found Courier superior to Times Roman in reading acuity but not in maximum reading speed, for readers with macular degeneration.

5.1. Width

The vast majority of fonts in books, newspapers, magazines, and web pages are "variable-width," which means that different letters may have different widths. For instance, letter 'i' is narrower than 'm'. In a variable-width font, average letter width is affected by whether capitals are included (they are usually wider than lowercase), whether widths are weighted by letter frequencies (narrow 't' is far more frequent than wide 'w'), and whether the "space" character is included (it is frequent and often narrower than the letters). In Times Roman, average lowercase width is 46% of body size, but average lowercase width including space character and weighted by character frequency is 41% of body.

Among variable-width typefaces for continuous reading at text sizes, the ratio of average lowercase letter width divided by x-height is slightly less than unity. For Times Roman: the ratio is 0.91; for Helvetica, 0.92.

In "fixed-width" ("monospaced") fonts like Courier, all letters and characters are the same width. Fixed-width Courier has a width/x-height ratio of 1.33, one-third greater than common variable-width fonts, because the font must accommodate structurally wide letters like 'm', 'w', 'M', and 'W', as well as structurally narrow letters like 'i' and 'l'. See Fig. 6.

Javal (1905) observed that average letter widths do not decrease in proportion to decrease of font body sizes. For example, the average width of a 6 point font is greater than half the width of a 12 point font. Partly anticipating Carter (1984/1937), Javal stated, "the engravers [punch-cutters] have recognized, without fully realizing it, that decrease of legibility is attributable principally to the decrease of width of the letters."

Arditi (1996) used the Metafont program to make test fonts of capital letters differing in width-to-height ratios. The fonts ranged from very narrow (width-to-height ratio = 1:5) to very wide (width-to-height ratio = 5:1). Arditi found that: "Fonts with width-to-height ratios less than unity are consistently more legible than their counterparts with reciprocal aspect ratio." Commonly used capitals, however, have width-to-height ratios very close to unity. For average capital widths: in Times Roman, width-to-height ratio is 1:1; in Helvetica, it is 1: 0.98; in Courier, 1: 1.04.

The widths of type letters include small white gaps on the left and right of the black letter forms (see dashed vertical lines in Fig. 3). These gaps or "sidebearings" determine inter-letter spacing and are set by type designers. Some word processing and typographic layout programs can

alter letter spacing by methods called “kerning,” “tracking,” or simply “spacing.”

Arditi, Knoblauch, and Grunwald (1990) made a fixed-width variation of Times Roman by adjusting inter-letter spacing so all letter widths were equal. They found that variable width fonts were read faster than fixed-width at large sizes but fixed-width fonts were read faster than variable-width at small sizes. In a true fixed-width font like Courier, the letter forms are designed to have equal widths, unlike a font in which spacing is altered but forms are not.

5.2. Weight (boldness)

Typeface weight is usually designated by impressionistic adjectives that may vary between typefaces and languages. The World Wide Web Consortium (W3C) defines font-weight on a numerical scale of 100 to 1000, from lightest to boldest, but the scale is imprecisely intuitive and ordinal, not ratio-based. Neither the names nor the W3C numbers are well-suited to psychophysical investigation.

Weight can be numerically measured by ink area (on screen, by black pixel percentages) or estimated by stem/x-height ratio. Pixel percentages can be measured with a photo editing program like Adobe Photoshop or GIMP, giving precise results with some complexity and effort. Type designers usually prefer to estimate weight easily by the ratio of vertical letter stem divided by x-height.

The black pixel percentage may be equal to or different from the stem/x-height ratio. For instance, in Times Roman, black pixels constitute 18% of a lowercase alphabet image, and the stem/x-height ratio is 1:5.5, equivalent to 0.18. In Times Bold weight, however, black pixels constitute 25% of the lowercase image, but the stem/x-height ratio is 1:3.3, equivalent to 0.30.

Bold weights mark text to signify emphasis or other structural distinction, and therefore are designed and selected to be distinctly different from normal text weight. In text type families, bold weights often range between 1.4 and 2.0 times normal weight, which seems to be an adequate distinction in common usage.

Many typeface families, especially sans-serif designs, now provide greater extremes of bold weights as well as of light weights. Weight differences can “fine-tune” text for perceptual or cultural reasons. When polarity is reversed – white type on black background – some typographers choose a bolder weight to prevent clogging of white letters in print or to make the text brighter on screen. In contexts related to fashion culture, light weights are often chosen for stylistic connotations, sometimes called “thin typography.”

In Fig. 7, eighteen weights of a single sans-serif type family are shown. The weight names are arbitrarily descriptive. The numbers approximate the W3C standard but are ratio-based. Weight 400 Normal is twice the weight of 200 Thin, based on stem/x-height ratio. The weight 800 Black is twice 400 Normal, and so on. Other type families may use different names, numbers, and scales for weights.

Luckiesh and Moss (1942) used a photometer to measure light reflected from pages printed in four different font weights – light, medium, bold, extra bold – of the typeface “Memphis.” They found that visibility (text viewed through filters) reached a plateau at the bold weight, while readability (measured by eye blink rate of readers) reached a plateau at the medium weight. Speed of reading reached a plateau between medium and bold weight, decreasing at the extra bold weight.

Paterson and Tinker (1940) measured speed of reading of normal and bold weight lower-case and found no difference, although 70 percent of the readers preferred the normal weight (Tinker, 1963).

In families of fixed-width fonts, weights can be compared by stem thicknesses. Bernard, Kumar, Junge, and Chung (2013) created six weights of Courier by adjusting stroke widths to 0.27, 0.72, 1.0, 1.48, 1.89, and 3.04 times the normal weight. The lightest font (weight = 0.27 × normal) and boldest (weight = 3.04 × normal) reduced reading speed of text viewed at the fovea. At 10 degrees eccentricity from the fovea, the two boldest weights (1.89 and

Weight Name	W3C	lowercase alphabet
UltraThin	100	abcdefghijklmnopqr
ExtraThin	150	abcdefghijklmnopqr
Thin	200	abcdefghijklmnopqr
ExtraLite	250	abcdefghijklmnopqr
Lite	300	abcdefghijklmnopqr
Book	350	abcdefghijklmnopqr
Text	375	abcdefghijklmnopqr
Normal	400	abcdefghijklmnopqr
Thick	425	abcdefghijklmnopq
ExtraThick	450	abcdefghijklmnopq
Dark	500	abcdefghijklmnopq
ExtraDark	550	abcdefghijklmnopq
Bold	600	abcdefghijklmnop
ExtraBold	650	abcdefghijklmnop
UltraBold	700	abcdefghijklmnop
Black	800	abcdefghijklmnop
ExtraBlack	900	abcdefghijklmnop
UltraBlack	1000	abcdefghijklmnop

Fig. 7. Weight progression of fonts of a sans-serif typeface family.

3.04 × normal) reduced reading speed. None of the experimental weights increased reading speed compared to normal weight.

5.3. Italic

The terms “roman” (upright) and “italic” (slanted) were borrowed into English from French words for Italian types of the early 16th century. Modern italic types may be slanted at angles roughly between 7 and 22 degrees, but common slants range from 10 to 16 degrees. Serifed italic faces are usually “cursive” (from Latin meaning “running”) because they imitate fast handwriting with different letter forms in addition to slant. “Oblique” styles, also called “slanted,” “inclined,” or “sloped” are slanted versions of roman, more common among sans-serif types, but see Fig. 8.

Roethlein (1912) found that recognition of italic letters was slightly lower than for corresponding roman letters.

Tinker (1963) found that lowercase italic type was read more slowly than roman by a small but significant amount that increased slightly during reading duration. Tinker also found that 96% of readers preferred roman type.

Wendt (1994) found a slight difference of 0.84 percent in reading speed favoring roman over italic.

Sheedy, Subbaram, Zimmerman, and Hayes (2005) found that, “Italics decreased the legibility of words and letters for block and letter size comparisons; however, only the effect on words was significant.”

It is not clear if the lower reading speeds and word recognition rates of italic types are because they are less legible in their features, or because they are unfamiliar to modern readers, except as a word or two amid roman text.

Roman abcdefghijklmnop
 Italic abcdefghijklmnop
 Oblique abcdefghijklmnop

Fig. 8. Roman, Italic, and Oblique of Times Roman.

5.4. Capitals vs. lowercase

Our capital letters come from inscriptional lettering of ancient Rome, but our lowercase letters come from Carolingian minuscule script of the Middle Ages. Italian Humanists combined capitals and minuscules into our familiar capital and lowercase alphabets, which were rendered into typography around 1465. In printers' jargon, "minuscule" became known as "lowercase" (or "lower-case") and capitals as "upper-case" because the different alphabets were held in different cases, the capital case above the minuscule case. Capital letters are usually taller and wider than the x-heights and widths of their lowercase counterparts and are approximately square in aspect ratio, following long tradition. Capitals constitute only 5% of average English text but have syntactic and semantic functions, such as marking sentence beginnings, distinguishing proper nouns from common nouns, and emphasizing text, so their legibility may result from interactions of function with form.

Tinker (1963) summarized studies by Paterson & Tinker and by Tinker, of the legibility of capitals and lowercase. Capitals were read more slowly than lowercase, by amounts ranging from 9.53 percent to 19.01 percent. Asked for preferences, 90 percent of readers preferred lowercase.

Poulton (1972) used word search and recognition methods to investigate legibility of capitals and lowercase in three typefaces at sizes near the threshold of legibility. In sans-serif "Univers" type, capitals were significantly less legible when equal to lowercase x-height. When adjusted so the legibility of capitals and lowercase were approximately equal, capital height was around 20 percent greater than lowercase x-height (1.44 mm and 1.19 mm, respectively).

Arditi and Cho (2007) compared reading speeds of readers with low vision or with normal vision, of sentences in all capitals, mixed-case, and all lowercase. Sentences were presented on-screen by RSVP and in continuous text. When font size was small (twice the acuity limit defined as the smallest size that could be read), capitals were read significantly faster by both kinds of readers in both kinds of presentations. When the font size was large (ten times the acuity limit) and presented to normally sighted readers, the advantage of capitals disappeared. The authors conclude that, "letter size determines legibility for low vision readers and for those viewing visually small text; and when point size is fixed, upper-case text is simply more legible...than lower-case." In Arial, the font in the study, capitals are approximately 184 percent greater in area than x-height lowercase.

5.5. Serifs

Serifs are small projections perpendicular to (or angled at) the terminals of main stems and strokes of letters. Sans-serif typefaces lack serifs. Times Roman is seriffed; Helvetica is sans-serif; see Fig. 6. Comparisons of seriffed to sans-serif typefaces have often found little or no significant difference in legibility.

Roethlein (1912) found that isolated letters of a sans-serif font, News Gothic, were more recognizable than those of fifteen seriffed fonts.

Paterson and Tinker (1932) measured reading speeds for ten fonts, including a sans-serif, Kabel Light, which was read slightly more slowly than six of the seriffed types, although the authors thought the difference barely significant. Tinker (1963) stated, "A serifless type, Kabel Light, is read as rapidly as ordinary type, but readers do not prefer it."

Zachrisson (1965) tested word recognition by school children of a seriffed versus a sans-serif font in tachistoscopic presentation and found no significant difference.

De Lange, Esterhuizen, and Beatty (1993) tested the legibility of Times Roman (seriffed) compared to Helvetica (sans-serif) fonts in texts scanned by primary school children searching for particular words. The font x-heights were adjusted to near equality, and line length and line spacing were equalized for the texts. The authors found no significant difference between seriffed and sans-serif.

Wendt (1994) compared legibility of sans-serif Futura with seriffed Bodoni and found slight but statistically non-significant superiority of the sans-serif. Futura and Bodoni, however, differ in several features other than serif and sans-serif.

Morris, Aquilante, Yager, and Bigelow (2002) tested legibility of seriffed and sans-serif fonts identical in body size, x-height, weight, width, contrast, phase (spacing of letter stems within the character cell), and inter-letter spacing, except that one font was seriffed and the other sans-serif. Using RSVP presentation, they found no difference in reading speed at a large size, equivalent to 16 point, but a 20% advantage for sans-serif at a very small size, equivalent to 4 point.

Arditi and Cho (2005) discuss claims that seriffed fonts are more legible than sans-serif, and used a specially designed parametric font with variants having serifs or no serifs, finding no significant differences in reading speed between the fonts in RSVP presentation or on paper.

Legge and Bigelow (2011) found that nearly all the typefaces in the running text of print newspapers and books were seriffed. More recently, on web sites, nine of the ten most used Google Fonts (around 900 free fonts served to web sites by Google) are sans-serif and constitute more than 70 percent of 30 trillion font downloads by Google (Google Fonts, 2019). The reasons for this apparent preference are not clear, but may be due to a belief that sans-serif fonts look "cleaner" on screens, or simply to the larger x-height fractions, averaging 0.52 of body size of the nine sans-serif Google fonts compared to smaller x-height seriffed fonts.

5.6. Typographic contrast

"Typographic contrast," shortened to "contrast" in typographic discourse, means the ratio of thickness of thick stem strokes to thin hairline strokes in letter forms. It is not luminance contrast as used in vision science. Among text typefaces: high-contrast faces like Didot or Bodoni have ratios between stem and hairline around 6:1; medium contrast faces like Baskerville and Times Roman have ratios around 4:1 to 4.5:1; low contrast faces like Jenson and Garamond have ratios around 2.5:1 to 3:1. Fig. 9 shows low, medium, and high contrast in seriffed text typefaces.

Sans-serif types like Helvetica and its look-alike Arial usually have low contrasts, around 1.2 to 1.5:1. Slab-serif types are similarly low contrast. Courier, a slab-serif, fixed-width face, has very low contrast, nearly 1:1, because its line thicknesses are mostly equal.

Typographic contrast has rarely been studied scientifically but provoked one of the most famous anecdotes in typographic history.

Benjamin Franklin was an admirer of English printer John Baskerville, designer of a new typeface with higher contrast than traditional Caslon types. In 1760, Franklin wrote to Baskerville about a "Connoisseur" who complained to Franklin that Baskerville's types would be the "Means of blinding all the Readers in the Nation, for the Strokes of your [Baskerville's] Letters being too thin and narrow, hurt the Eye, and he [Connoisseur] could never read a line of them without Pain." Franklin removed the name from a Caslon type specimen and



Fig. 9. Typographic Contrast: low (Jenson), medium (Baskerville), high (Didot).

tricked the Connoisseur into thinking it came from Baskerville, whereupon the Connoisseur pointed out the “Disproportion” of the type and complained he could not read it without feeling pain, not noticing he was criticizing his favorite Caslon type, not Baskerville’s. (Franklin, 1760). Modern versions of Baskerville types remain popular in academic publishing today.

Paterson and Tinker (1932) compared reading speeds of seven common text typefaces plus three less common faces. Bodoni, a high-contrast font, was read 1.4% more slowly than Garamond, an Old Style font with lower contrast, but the reading speed difference was not statistically significant.

Wendt (1994) found statistically non-significant legibility difference between low-contrast Futura and high-contrast Bodoni typefaces.

Mansfield et al. (1996) found that low contrast Courier was read faster than medium–high contrast Times at a small size near the acuity limit, but did not determine whether the greater legibility of Courier at the small size was attributable to the greater letter widths of Courier or to its lower contrast.

Larson and Carter (2016) found that the size-specific variant of the Sitka typeface with lowest contrast was more legible at both large and small sizes than the variant with higher contrast. There were differences of other features, including weight, width, and x-height, so it is not clear to what extent contrast alone affected legibility.

5.7. Spacing

“Spacing” generally refers to space between letters. Some word processing or typographic layout programs can “kern,” “track,” “space,” “expand” or “condense” letter spacing. However, typefaces specifically named “expanded” or “condensed” are designed to be wider or narrower than normal width, not simply spaced differently. Here we cite a few of many studies of letter spacings.

Huey (1908), reviewing earlier studies, stated: “There is probably little to be gained by increasing the distance between the letters beyond that which is usual in the better printed books of the present time.... It should be remembered that any very unusual separation of the letters of a word is distracting and should be avoided.”

Chung (2002) tested the effects of different letter spacings on reading speed in central and peripheral vision by adjusting letters in the fixed-width Courier font through a range of five spacings: half normal; 0.707 normal; normal (original); 1.414 normal; 2.0 normal. These were tested at visual sizes above and below critical print size. The results revealed a “critical letter spacing,” above which reading speed became constant as spacing increased, until it decreased slightly at much greater distances. Below critical spacing, reading speed decreased.

In studies of dyslexia, Spinelli, De Luca, Judica, and Zoccolotti (2002), Zorzi et al. (2012), as well as Perea, Panadero, Moret-Tatay, and Gómez (2012) reported benefits of increased letter spacing in accuracy, speed, or word identification by dyslexic readers. However, Schneps et al. (2013) found that increased letter spacing offered only mixed benefits to struggling readers, but that shorter text lines, in number of characters, offered clear benefits including dramatic increase in reading speed.

6. Types designed with reference to legibility studies

Despite early optimism that scientific study of reading would improve legibility of type, few commercial typefaces have been based on legibility research.

Javal (1905) showed a sample word of an experimental typeface designed by his associate Charles Dreyfuss, using Javal’s ideas about legibility for very small type sizes. See Fig. 10.

The Javal-Dreyfus example was not a commercial font, but 101 years later, a digital typeface for very small sizes, named “Minuscule,” was designed by Thomas Huot-Marchand, influenced by, although not exactly like, the Javal-Dreyfuss example. See Fig. 11.



Fig. 10. Type design for very small sizes, by Charles Dreyfuss and Émile Javal, 1905.



Fig. 11. Minuscule typeface for very small sizes, by Thomas Huot-Marchand, 2006.

6.1. Century Schoolbook 1920

In 1915, textbook publisher Ginn & Company asked American Type Founders (ATF) to develop a typeface to benefit the reading and eyesight of children. The result, released in 1920, was Century Schoolbook, designed by Morris Fuller Benton, who had consulted with Barbara Roethlein on her Clark University legibility thesis (Roethlein, 1912). Benton may also have heard of Javal’s research from Theodore Low De Vinne, an eminent printer-publisher who corresponded with Javal and had worked with Benton at ATF (Cost, 2011). Century Schoolbook is still current, now in digital fonts formats a century after its first release.

In promoting Century Schoolbook, ATF (1923) stated: “A series of tests and experiments made by the British Association for the Advancement of Science, Clark University, the Board of Education of New York City, the American Type Founders Company and others, all go to show that some styles of type are far more legible than others.” (see Fig. 12)

6.2. Linotype legibility group

In the 1920s, the Linotype corporation, which manufactured the majority of typesetting machines used by American newspapers, began to develop space-saving types to reduce newsprint costs while maintaining legibility. The Linotype “legibility group” eventually included Ionic No. 5 (1925), Excelsior (1931), Opticon (1936), Paragon (1936), and Corona (1941). By the early 1950s, Corona was one of the most widely used typefaces in American newspapers. Its x-height was 20% greater than the first of the legibility group, Ionic No. 5.

Linotype stated: “Though Ionic No. 5 appears to be much larger than the average newspaper face, and is considerably easier to read, it actually takes up no more space in newspaper and periodical columns. Thus it does not minimize letter and word count, and is an excellent choice where maximum matter per page, and the desirable factor of reading ease, are required. [It] was originally designed to overcome the handicap of eye fatigue and impaired vision of many newspaper readers, particularly those of advanced years. Its letter design is based upon a study of eye movements in reading, as well as on the practical requirements of modern printing and stereotyping methods.” (Linotype Specimen Book, n.d., circa 1940.)

The Linotype Legibility Group typefaces, for instance, Excelsior, resemble Century Schoolbook, which may have been a unacknowledged model for the Linotype “legibility” series. See Fig. 13.

Century Schoolbook



Fig. 12. Century Schoolbook, by Morris Fuller Benton, 1920.

Excelsior
 H a b c d e f g h i j k l m n o p q r s

Fig. 13. Excelsior, by Chauncey Griffith and Linotype design team, 1931.

6.3. Times New Roman, 1931

Stanley Morison, in his 1930 “Memorandum on a Proposal to Revise the Typography of ‘The Times’” (Morison, 1980), discussed a finding by Pyke (1926) that some Old Style types were more legible than Modern style types. (“Old Style” refers to types resembling styles cut before 1700, and “Modern” to types resembling styles cut after 1800.) “Times New Roman,” designed in 1931 by Morison and Victor Lardent, resembles an Old Style typeface cut circa 1570 but has a larger x-height fraction and narrower lowercase width (Dreyfus, 1973). See Fig. 6.

Vision scientists may wish to know if fonts of “Times New Roman” in TrueType format and “Times Roman” in PostScript format are identical. Very nearly so. Their character widths relative to body size are essentially the same, but in x-height fraction, Times Roman is very slightly greater than in Times New Roman: 0.450 versus 0.447, respectively. In vertical lowercase stems, those of Times Roman are very slightly thicker than those of Times New Roman: 0.084 of body versus 0.081, respectively. These differences may be negligible in most reading contexts. At 12 point size on a resolution of 300 pixels per inch, the respective x-heights are less than a pixel different, and the stem thicknesses less than 0.15 pixel different.

7. Digital types designed with legibility studies

The emergence of digital typography in the 1980s prompted development of fonts for computer screens and laser printing. Most early digital fonts were digitizations of traditional analog designs, but a few designers as well as reading researchers consulted relevant research.

7.1. Lucida

In 1984–1985, Charles Bigelow and Kris Holmes designed “Lucida” (Bigelow & Holmes, 1986; 2018), a family of typefaces partly influenced by vision and reading research. Lucida’s large x-height and wide inter-letter spacing were influenced by Bouma (1970) on crowding, by Campbell and Robson (1968) on spatial frequency, and by Tinker (1963) on factors of legibility. Lucida fonts have been widely used in screen-based computer-user interfaces (Macintosh OS X, Bell Labs Plan 9, Microsoft Windows. See Fig. 14.

7.2. Font Tailor

“Font Tailor” is an interactive tool developed by Arditi (2004) that enabled readers with low vision to adjust parameters of a basic font until it appeared most legible. Although the visually impaired users produced a variety of distinct fonts, the study did not find the parametrically adjusted fonts to be more legible than Times Roman.

7.3. Tiresias

“Tiresias PCfont” is a sans-serif typeface with a large x-height fraction commissioned by the Royal National Institute of Blind People

Lucida Grande
 H a b c d e f g h i j k l m n o p q r s

Fig. 14. Lucida Grande for medium resolutions, by Bigelow & Holmes, 2001.

for readers with impaired vision. Rubin, Feely, Perera, Ekstrom, and Williamson (2006) compared reading speeds for Tiresias PCfont to Times New Roman, Helvetica, and “Foundry Form Sans,” and found that when the size of Tiresias PCfont was adjusted to equalize its actual vertical and horizontal space to the space occupied by the other fonts, Tiresias PCfont did not provide an advantage in reading speed.

7.4. OvinkTest, PykeTest, SpencerTest, and Neutral Test

“OvinkTest,” “PykeTest,” and “SpencerTest” fonts were designed by Sofie Beier in 2009 for research on legibility. For each typeface, Beier and Larson (2010) tested recognition of different versions of frequently misrecognized letters, using short exposure to the parafovea at a distance of 50 cm, and a separate recognition test at an average distance of 6 m. The study did not find significant differences between some variant letter designs, but recognition of narrow letters: f, j, l, t was greater for the wider versions of the letters. Beier later modified the OvinkTest font and released it as a commercial font named “Ovink.”

Neutral Test (Beier, 2013) is a type family for legibility testing. Each font in the family differs from the others in one or two distinct features, such as letter skeleton, weight, width, or contrast. Dyson and Beier (2016) used Neutral Test to find that italic style is more subtle than bold when used for emphasis.

7.5. Afandem Arabic

Nadine Chahine (2012) designed three variants of an Arabic typeface in Naskh style: Afandem Traditional is a style typical for book composition; Afandem Simplified omits some context-sensitive Arabic letter variations; Afandem Dynamic includes variant letterforms, joined letter combinations, and stackable alternate forms to simulate Arabic calligraphy. Eye-tracking of students reading texts in the variants of Afandem found differences in fixation duration – shorter for Simplified and longer for Dynamic – but no significant differences in overall reading speeds. Concerning fixation duration, one conclusion was that, “increased complexity of word formation has a negative effect on the legibility of Arabic typefaces.”

7.6. Matilda

Ann Bessemans designed the “Matilda” typeface family following Ph.D. research on font design for children with impaired vision (Bessemans, 2012). Bessemans studied the effects of modifications of features in commercial serifed and sans-serif typefaces, by testing the fonts with school children with normal vision and with low vision. Based on the findings, she designed new typefaces to support early stage reading for visually impaired schoolchildren, and Matilda was a main outcome (Bessemans, 2016).

7.7. Sitka

Larson and Carter (2016) tested legibility of “size-specific” (“optically scaled”) variants of the “Sitka” typeface family designed by Matthew Carter. (See Fig. 5.) During the design process, the authors used letter recognition tests to compare legibility of the smallest size-specific design (Sitka Small) to the largest size-specific design (Sitka Banner), making adjustments to improve recognition. In Fig. 15, Sitka Small has a substantially larger x-height (0.54 of body) than Sitka Banner (0.47 of body) and is approximately 30% wider. The authors found that the “Small” version performed better at both small and large sizes, concluding that: “This strongly indicates that the size-specific adjustments made for large sizes do not increase legibility for large sized text. If we want increased legibility at large sizes, we are better served using a small size-specific design.”

Sitka Banner

Habcde fghijklmnopqrstuvw x

Sitka Small

Habcde fghijklmnopqr

Fig. 15. Sitka Small and Sitka Banner, by Matthew Carter, 2016.

7.8. Eido and Maxular Rx

Two fonts intended to improve reading for readers with macular degeneration are “Eido” (Bernard, Aguilar, & Castet, 2016) and “Maxular Rx” by Steven Skaggs. (Fig. 16.) Xiong, Lorsung, Mansfield, Bigelow, and Legge (2018) compared these two fonts to three widely used typefaces – Times Roman, Helvetica, and Courier. Eido lowercase letters differ from each other in features of weight, slant, x-height, and cursiveness, to achieve greater differentiation of letters. Maxular Rx is consistent in featural parameters but is bolder and wider and has lower typographic contrast than Eido, and than Times, Helvetica, or Courier. See Fig. 16. The authors summarize their findings thus:

“Compared to Helvetica and Times, Maxular Rx permitted both smaller CPS [critical print size] and RA [reading acuity], and Eido permitted smaller RA. However, the two new fonts did not present any advantage over Courier. Font variations had less influence on the reading performance of normally sighted subjects, affecting only RA.”

8. New forms of fonts

Digital print and screen types have tended to follow the black & white traditions and forms of writing and printing, but high-resolution color displays on mobile phones, computer screens, and other digital reading media may now include color, gray tone, motion, and other features on which non-traditional type forms may be based. Moreover, popular emergence of emoji pictograms and ideograms with colors, tones, and intricate details is a proliferating phenomenon that likewise invites study.

Experimental fonts for vision research are often optotypes containing only a few letters or symbols designed to test thresholds of perception and recognition of letter and number forms. These do not follow the esthetics expressed by Bodoni (1818), or the constraints of economy and tradition that Tinker (1963) believed had slowed development of scientific typography. Nevertheless, new experimental fonts advance the study of legibility in directions not anticipated in earlier research. Several new optotypes have come from the Pelli Lab.

8.1. Sine wave modulated letters

Oruç, Landy, and Pelli (2006) created “second-order” gray-tone letters to investigate spatial frequency channels in letter recognition. Sloan optotype letters were rendered as modulated vertical and horizontal sine wave patterns, as seen in Fig. 17. The modulated letters lack high luminance contrast but can be recognized. They cannot be rendered with standard digital font technology, but can be displayed and printed by other imaging techniques, and while they may not conform

Eido

HABCDEFghijklmno pqrstUvwxyz

Maxular

Habcde fghijklmnopqrstu vwxyz

to Bodoni’s express criteria for beauty, they have enough spooky charm to recommend the creation of a full alphabet.

8.2. Checkers alphabets and perimeteric complexity

Pelli, Burns, Farell, and Moore-Page (2006) created simple “Checkers” alphabets to investigate character complexity and feature detection in letter recognition. The alphabets comprise 26 simple, random block patterns of 2×3 and, separately, 4×4 block patterns. The authors state that the 2×3 Checkers font is analogous to Braille and can be learned and read efficiently after a sufficient number of trials. This paper also introduces and analyzes a feature not previously described in typography: “perimeteric complexity,” which is expressed by the formula: perimeter squared divided by interior area times 4π . In laboratory study, Pelli et al. found that efficiency of letter recognition is greater for simpler forms than for complex. Typographers appear to agree implicitly with this finding, by usually using types with simpler letter forms for running text at small sizes, but occasionally using types of greater formal complexity for brief texts at large sizes. Watson (2012) discusses the mathematics of perimeteric complexity in geometric and digital images, inadvertently suggesting why perimeteric complexity was not analyzed in traditional typography. Nevertheless, some of the finest artists among traditional type punchcutters, including Robert Granjon, Pierre-Simon Fournier, and Bodoni himself when young, cut entrancingly complex ornamental characters that were assembled into patterns but not alphabets.

8.3. Pelli font

In developing a clinical test for visual crowding, a limitation on recognition of cluttered symbols, Pelli et al. (2016) designed an optotype font of numerals with an extremely narrow aspect ratio of 5:1 – tall but narrow – which can be identified at tiny widths. Fig. 18 compares the numerals of the Pelli font to the letters of the familiar Sloan optotype font.

9. Smartfonts & Livefonts

Bragg, Azenkot, and Kalai (2016) created novel “Smartfonts” as block pixelated color fonts, which, the authors state, “completely redesign the written alphabet with the purpose of improving the reading experience.” Smartfonts require readers to learn to read the unfamiliar symbols, which may prompt new ways of measuring learnability and legibility. After Smartfonts, Bragg, Azenkot, Larson, Bessemans, and Kalai (2017) introduced “Livefonts,” which incorporate animated motion as well as color and simple geometric shapes, to possibly enhance reading with low vision. The characters of Livefonts are colored shapes that may pulse, flash, jump, or rotate at various rates. These new kinds of fonts do not yet demonstrate marked improvements over traditional type, but are inventive explorations of electronic media for typographic display.

10. Conclusion

The first century of legibility research focused on practical effects of typeface features on reading and publishing economy. In recent

Fig. 16. Eido by Bernard et al. and Maxular RX by Steven Skaggs.

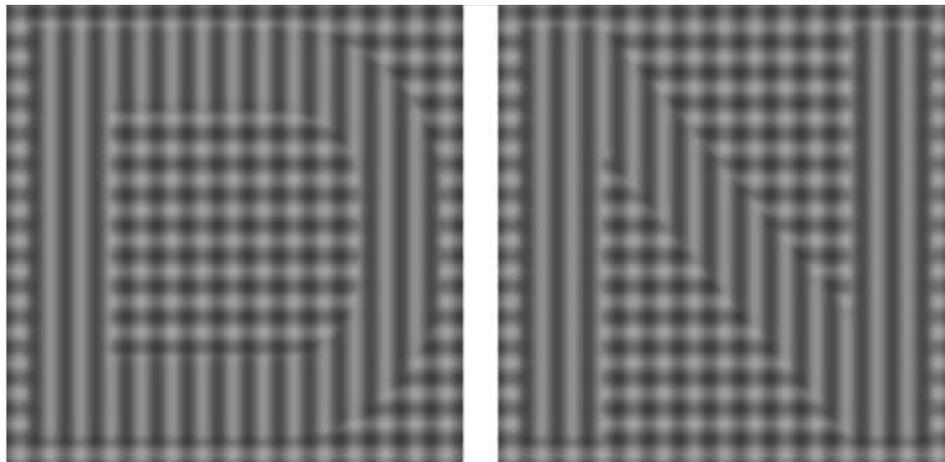


Fig. 17. Sine wave modulated Sloan letters D N.



Fig. 18. Pelli font numerals compared to Sloan letters.

decades, digital typographic tools have provided more precision along with greater ease of experimentation. Following on earlier studies, digital type techniques have led to new insights and broader agreement on the effects of type size on reading, and to some agreement on the effects of width and weight. Some perennial questions of comparative legibility, such as seriffed versus sans-serif types, or capitals versus lowercase, have continued to resist easy analysis. In general, digital type technology has greatly facilitated general studies of the psychophysics of reading and legibility, and in particular the design and testing of fonts for low vision, macular degeneration, developmental dyslexia, and other reading difficulties neglected by commercial typography. Most recently, novel fonts of surprising forms have been created for fundamental investigations of visual phenomena including crowding, perimetric complexity, and spatial frequency channels. The current florescence of innovative fonts for digital reading platforms may perhaps justify the optimism expressed by early legibility researchers more than a century ago.

Acknowledgements

The author wishes to thank: Kris Holmes for preparation of Figs. 2, 4 & 7; Monotype Imaging for permission to use Fig. 1; Denis Pelli for permission to use Figs. 17 and 18; Gordon E. Legge for helpful comments and questions on an early draft; and the VR editors and reviewers for cogent comments and critiques. The author is solely responsible for any errors of fact or interpretation.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Arditi, A. (1996). Typography, print legibility, and low vision. *Remediation and Management of Low Vision*, 1, 237–248.
- Arditi, A. (2004). Adjustable typography: An approach to enhancing low vision text accessibility. *Ergonomics*, 47(5), 469–482.
- Arditi, A., & Cho, J. (2005). Serifs and font legibility. *Vision Research*, 45(23), 2926–2933.
- Arditi, A., & Cho, J. (2007). Letter case and text legibility in normal and low vision. *Vision Research*, 47(19), 2499–2505.
- British Association for the Advancement of Science. (1913). Report on the Influence of School-Books upon Eyesight. Offices of the Association.
- Beier, S. (2012). *Reading letters: Designing for legibility*. BIS Publishers.
- Beier, S., & Larson, K. (2010). Design improvements for frequently misrecognized letters 1. *Information Design Journal*, 18(2), 118–137.
- Beier, S. (2013). Legibility investigation: Towards controlling typeface variables. Praxis and Poetics: Research Through Design, Conference, Newcastle.
- Bernard, J. B., Aguilar, C., & Castet, E. (2016). A new font, specifically designed for peripheral vision, improves peripheral letter and word recognition, but not eye-mediated reading performance. *PLoS One*, 11(4), e0152506.
- Bernard, J. B., Kumar, G., Junge, J., & Chung, S. T. (2013). The effect of letter-stroke boldness on reading speed in central and peripheral vision. *Vision Research*, 84, 33–42.
- Bessemans, A. (2016). Matilda: A typeface for children with low vision. *Digital fonts and reading* (pp. 19–36). World Scientific.
- Bigelow, C., & Holmes, K. (1986). *The design of Lucida. Text processing and document manipulation*. Cambridge University Press.
- Bigelow, C., & Holmes, K. (2018). Science and history behind the design of Lucida. *TUGboat*, 39(3), 204–211.
- Bigelow, C. (2016). Lisibilité et typographie: les recherches durant la première moitié du XXe siècle. In Histoire de l'Écriture Typographique: le XXe siècle de 1900 à 1950, Jacques André, ed. Atelier Perrousseau (pp. 194–213).
- Bodoni, G. (1818). *Manuale Tipografico*. Presso la Vedova. (pp. XXI–XXXVIII). (Facsimile reprint (2010), S. Füssel, ed. Taschen.).
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226(5241), 177.
- Bragg, D., Azenkot, S., & Kalai, A. T. (2016, October). Reading and learning smartfonts. In Proceedings of the 29th annual symposium on user interface software and technology. ACM (pp. 391–402).
- Bragg, D., Azenkot, S., Larson, K., Bessemans, A., & Kalai, A. T. (2017, October). Designing and evaluating livefonts. In Proceedings of the 30th annual ACM symposium on user interface software and technology. ACM (pp. 481–492).
- Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology*, 197(3), 551–566.
- Carter, H. (1984). Optical scale in type founding. *Printing Historical Society Bulletin*, 13, 144–148. (Reprint from Carter, H. (1937). Optical scale in typefounding. *Typography* 4.
- Chahine, N. (2012). Reading Arabic: Legibility studies for the Arabic script (Doctoral dissertation, Faculty of the Humanities, Leiden University).
- Chung, S. T. (2002). The effect of letter spacing on reading speed in central and peripheral vision. *Investigative Ophthalmology & Visual Science*, 43(4), 1270–1276.
- Cost, P. A. (2011). *The Bentons: How an American Father and son changed the printing industry*. RIT Press.
- CSS (<https://www.w3.org/wiki/CSS/Properties/font-weight>).
- De Lange, R. W., Esterhuizen, H. L., & Beatty, D. (1993). Performance differences between Times and Helvetica in a reading task. *Electronic Publishing*, 6(3), 241–248.
- Dreyfus, J. (1973). The evolution of Times New Roman. *Penrose Annual*, 66, 165–174.
- Dreyfus, J. (1995). The Baskerville Punches. In *Into Print: Selected Writings on printing history, typography and book production*. David R. Godine, Boston. pp. 13–36.
- Dyson, M. C., & Beier, S. (2016). Investigating typographic differentiation: Italics are more subtle than bold for emphasis. *Information Design Journal*, 22(1), 3–18.

- Font Tools urls (2019). (accessed April 29, 2019) Fontforge: <https://fontforge.github.io/en-US/> FontLab, Fontographer, TypeTool, OTMaster: <https://www.fontlab.com/Glyphs/> <https://glyphsapp.com>.
- Franklin, B. (1760). Letter to John Baskerville. <https://founders.archives.gov/documents/Franklin/01-09-02-0085> (accessed April 29, 2019).
- Gable, N. & Paput, C. (2016). Pérennité des poinçons et matrices. In *Histoire de l'Écriture Typographique: le XXe siècle de 1900 à 1950*, Jacques André, ed. Atelier Perrousseaux (pp. 88–113).
- Griffing, H., & Franz, S. I. (1896). On the condition of fatigue in reading. *The Psychological Review*, 3(5), 513–530.
- Huey, E. B. (1908). *The psychology and pedagogy of reading*. The Macmillan Company.
- Javal, E. (1905). *Physiologie de la lecture et de l'écriture*. Paris: Félix Alcan.
- Knuth, D. E. (1986). *Computers and Typesetting, Volume C: The Metafont Book*. Addison-Wesley.
- Larson, K., & Carter, M. (2016). Sitka: A collaboration between type design and science. *Digital Fonts and Reading* (pp. 37–53). World Scientific.
- Legge, G. E. (2006). *Psychophysics of reading in normal and low vision*. Lawrence Erlbaum Associates.
- Luckiesh, M., & Moss, F. K. (1942). Reading as a visual task. D. Van Nostrand (pp. 165–178).
- Lund, O. (1999). *Knowledge construction in typography: The case of legibility research and the legibility of sans serif typefaces (Doctoral dissertation)*. University of Reading.
- Mansfield, J. S., Legge, G. E., & Bane, M. C. (1996). Psychophysics of reading. XV: Font effects in normal and low vision. *Investigative Ophthalmology & Visual Science*, 37(8), 1492–1501.
- Mansfield, J. S., Legge, G. E., Luebker, A., & Cunningham, K. (2006). The MNREAD acuity chart. *Psychophysics of Reading in Normal and Low Vision*, 167–191.
- Morison, S. (1980). Memorandum on a Proposal to Revise the Typography of 'The Times'. In *Selected essays on the history of letter-forms in manuscript and print*. David McKitterick, ed. Vol. 2. Cambridge University Press (pp. 295–317).
- Morris, R. A., Aquilante, K., Yager, D., & Bigelow, C. (2002). P-13: Serifs Slow RSVP Reading at Very Small Sizes, but Don't Matter at Larger Sizes. In *SID Symposium Digest of Technical Papers* (Vol. 33, No. 1, pp. 244–247). Blackwell Publishing Ltd.
- Oruc, I., Landy, M. S., & Pelli, D. G. (2006). Noise masking reveals channels for second-order letters. *Vision Research*, 46(8–9), 1493–1506.
- Ovink, G. W. (1938). Legibility, Atmosphere Value and Forms of Printing Types. A.W. Sijthoff's Uitgeversmaatschappij N.V. Leiden.
- Paterson, D. G., & Tinker, M. A. (1932). Studies of typographical factors influencing speed of reading. X. Style of type face. *Journal of Applied Psychology*, 16(6), 605.
- Paterson, D. G., & Tinker, M. A. (1940). *How to make type readable*. Harper & Brothers Publishers.
- Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, 46(28), 4646–4674.
- Pelli, D. G., Waugh, S. J., Martelli, M., Crutch, S. J., Primativo, S., Yong, K. X., ... Yiltiz, H. (2016). A clinical test for visual crowding. *F1000Research*, 5.
- Perea, M., Panadero, V., Moret-Tatay, C., & Gómez, P. (2012). The effects of inter-letter spacing in visual-word recognition: Evidence with young normal readers and developmental dyslexics. *Learning and Instruction*, 22(6), 420–430.
- Poulton, E. C. (1972). Size, style, and vertical spacing in the legibility of small typefaces. *Journal of Applied Psychology*, 56(2), 156.
- Pyke, R. L. (1926). Report on the legibility of print. Medical Research Council, Special report series, no. 110. His Majesty's Stationery Office: London.
- Roethlein, B. E. (1912). The relative legibility of different faces of printing types. *The American Journal of Psychology*, 23(1), 1–36.
- Rubin, G. S., Feely, M., Perera, S., Ekstrom, K., & 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*, 26(6), 545–554.
- Schneps, M. H., Thomson, J. M., Sonnett, G., Pomplun, M., Chen, C., & Heffner-Wong, A. (2013). Shorter lines facilitate reading in those who struggle. *PLoS One*, 8(8), e71161.
- Sheedy, J. E., Subbaram, M. V., Zimmerman, A. B., & Hayes, J. R. (2005). Text legibility and the letter superiority effect. *Human factors*, 47(4), 797–815.
- Spinelli, D., De Luca, M., Judica, A., & Zoccolotti, P. (2002). Crowding effects on word identification in developmental dyslexia. *Cortex*, 38(2), 179–200.
- Tinker, M. A. (1963). *Legibility of print*. Iowa state University Press.
- Watson, A. B. (2012). Perimetric complexity of binary digital images: Notes on calculation and relation to visual complexity. *The Mathematica Journal*, 14.
- Wendt, D. (1994). What do we mean by legibility. In Peter Karow (Ed.). *Font Technology* (pp. 271–306). Springer Verlag.
- Xiong, Y. -Z., Lorsung, E. A., Mansfield, J. S., Bigelow, C., & Legge, G. E. (2018). Fonts Designed for Macular Degeneration: Impact on Reading.
- Zachrisson, B. (1965). *Legibility of printed text*. Almqvist & Wiksell.
- Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M., ... Ziegler, J. C. (2012). Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences*, 109(28), 11455–11459.