Transforming our World Through Design, Diversity and Education
G. Craddock et al. (Eds.)

doi:10.3233/978-1-61499-923-2-862

Legibility in Print Text for People with Impaired Vision

Jonny NERSVEEN[a,1], Anne Kristin KVITLE[a] and Eivind Arnstein JOHANSEN[b]
[a] *Department of Manufacturing and Civil Engineering, NTNU Norwegian University of Science and Technology*
[b] *Department of Design, NTNU Norwegian University of Science and Technology*

Abstract. Research on the level of legibility and readability of text are mainly based on subjects with normal eyesight. In Norway, about 180.000 of the population is diagnosed with a visual impairment. In our study, over 800 visually impaired subjects participated in an experiment, which is so far the largest study of legibility and readability ever conducted for this this group of observers. The observers were recruited through the Norwegian association for blind and visually impaired (Norges Blindeforbund) and the number of subjects reveals that the experiment included 4.6 % of the population with visual impairments. In the experiment, the characteristics to be studied included different typefaces, serifs and sans serifs, font sizes, weighting and contrast. We can conclude that the uncontrolled home based experiment was successful regarding the volume of respondents, which resulted in significant findings. This paper gives an overview of the design choices according to experiment method and selection of legibility variables.

Keywords. Impaired vision, legibility, readability

1. Introduction

The idea of this study was conceived when the Norwegian Research Laboratory for Universal Design was contacted by the Delta Centre of Norway, the national resource center for accessibility and social inclusion, with a request of how the Delta Centre should provide accessibility and universal design when publishing printed material. People with dyslexia and visual impairments were among the target groups for enabling accessibility, and a literature review on readability and legibility in printed text for such observers was conducted as a part of the request from the Delta Centre. This review [1] revealed two findings in existing research concerning partially sighted people: either the studies had a small number of observers, or the studies were limited to target a specific visual impairment. For this reason, existing results could not be generalized to evaluate what parameters of print text guide legibility for the partially sighted. Consequently, there was a need to collect new data which involved designing and conducting an experiment including several different visual impairments. As the condition of being partially sighted may include several different visually impairments (most observers had more than one), the aim was to cover the most common types. Also, it was a need to

[1] Corresponding Author, The Norwegian Research Laboratory for Universal Design, NTNU Gjøvik, Teknologivegen 22, 2815 Gjøvik, Norway; E-mail: jonny.nersveen@ntnu.no.

examine several characteristics of the printed text in order to evaluate the legibility. As the factors involved (visually impairments and parameters of text) were wide-reaching, there was a need for a large number of participants.

As the terms of media (print) and observer group (partially sighted) were decided, some of the main challenges in this study was to select the appropriate text parameters to examine and to recruit a large number of observers. The aim of this paper is to describe the main considerations and decisions in the experiment design that resulted in one of the largest studies of legibility for partially sighted observers conducted.

2. Background

Readability and legibility are terms that have sometimes been, and perhaps still are, used more or less interchangeable [2]. Readability is the property that enables reading and applies to the reading experience. Legibility is the visual properties of the text, which is critical for an efficient readability. Legibility includes basic characteristics of typeface design [3] such as serifs, font sizes, weight, contrast and more, and are studied in previous research [4-7].

A typical task in visual experiments is for the subjects to perform a visual task like match colors or to read a text, or to evaluate something according to preference or rank. Visual appearance depends on viewing parameters such as illumination and surround conditions and viewing geometry. For this reason, visual experiments are traditionally carried out as controlled experiments in a laboratory where the experimental setup and laboratory settings are calibrated, controlled and measured according to requirements, best practice or standards. However, such experiments are both time consuming and expensive to carry out, and often limited to a small number of observers. Uncontrolled experiments can be carried out in the field or on the web. Some of the major advantages of the uncontrolled experiments are that they can be easy to carry out, the observer can do the experiment in their own home and it has a potential of reaching and recruiting a large number of observers.

Uncontrolled experiments can be web based, where display technology and software are among a large number of uncontrolled variables. For print media, the light sources are similar uncontrolled variables. However, these uncontrolled experiments include real-world sources of variability [8] which is both an advantage and dis-advantage. Another argument [9] is to what extent a controlled experiment is representative of a real world and suggest that the different viewing conditions could be regarded as “a source of noise in the recorded data, whose could be compensated by a high number of participants, producing even more significant results, not biased by a specific set of viewing conditions”.

Research comparing the controlled and uncontrolled experiment methods supports the hypothesis of equivalence between the methods, concluding that the uncontrolled can substitute the controlled experiment when the number of observers is large [8-10].

Previous studies in evaluation of prints [10], where the same source material where evaluated in controlled and uncontrolled viewing conditions, show equivalence between the methods.

As a study of legibility would demand a great number of varieties of visual impairments among the observers, consequently a large number of observers were required. For that reason, an uncontrolled experimental method was chosen. To reach a significant number of participants, the study was designed as a home based experiment

where the observers received a booklet with the text to be studied and a questionnaire. The experiment is described in the next section.

3. Method

The aim of the study was to examine and verify which parameters in printed text that guided readability for visually impaired observers, and for some parameters (i.e. size) evaluate the thresholds for readability. To achieve this, subjective assessment is required where human observers are asked to grade printed text according to a given set of legibility criteria. In our study, the design choices can be categorized according to

- Subjects
- Legibility parameters
- Experimental method

and are discussed and described in the following sections.

3.1. Subjects

The observers were recruited among the members of the Norges Blindeforbund, and the booklet including the text and questionnaire was distributed to a test population of 5.000 members (of a total of approximately 9.700). Since the observers may suffer from several eye diseases, the observers were asked to specify their impairments based on the following selection list:

- Corneal diseases
- Aniridia
- Albinism
- Iridocyclitis
- Retinis pigmentosa
- Cataract
- Glaucoma
- Age related macular degeneration
- Juvenile macular degeneration
- Diabetic retinopathy
- Optic nerve disorders
- Scotoma caused by stroke
- Other visual impairments caused by stroke
- Other visual impairments

The participants are all 18 and older, male and females. To qualify as a test person, one must be categorized as partially sighed and be able to read printed text with only the aids allowed in this study (magnifying glass, magnifying lamp, supplementary lights etc.). All participants were randomly selected.

3.2. Legibility parameters

The aim of the study was to examine which parameters in printed text that guided readability, and for some parameters (i.e. size) evaluate the thresholds for legibility. The parameters to be examined were typeface, font size, serif or sans serif, font weight and positive and negative contrasts. One of the main challenges in this part of design phase was to limit the number of parameters so the tasks in the experiment would not be overwhelming to the subject. Other parameters, such as color and color contrast, were considered but omitted due to restrictions of the size and complexity of the study.

3.2.1. Characteristics of typefaces

Composition of a typeface and its characteristics is a science of its own, but a short introduction of characteristic [3] is needed to explain and justify the decisions made in designing the experiment. Each letter in a typeface has a distinct width and a height, where the lowercase letter height is often referred to as the "x height" (see figure 1). This is the defined distance between the baseline and the mean line of lower-case letters in a typeface. The height of a typeface, also referred to as body height or kegel, defines the height of the typeface, as a minimum this equals generally the distance between the ascender and the descender, as illustrated in figure 1.

Figure 1. X height and body height in typeface.

A parameter in previous research is the relation between the x height and the font size [4]. The x height may vary for different typefaces in the same font size. This is illustrated below, where three different typefaces are shown in font size 12:

Verdana font size 12 pt

Arial font size 12 pt

Times New Roman font size 12 pt

For the typefaces and sizes to be comparable in the experiment, the x height of each of the typefaces needed to be scaled. The typeface Times roman is used as a scaling reference, where the x height of each typeface is scaled to match the x height of Times roman.

Serifs have been in use for centuries [5] and the debate whether serifs enable readability is an ongoing debate [6][7][8]. In print, Times is a frequent used typeface with serifs. Tiresias is a typeface designed to aid readability for partly sighted persons, and is without serifs. The guidelines from Norges Blindeforbund recommend sans serifs for better readability for partly sighted persons. The literature review [1] did not reveal any research proving increased readability with sans serif typefaces.

Another characteristic of a typeface is emphasis, where several techniques are used such as **boldface** or *italics*. A faux boldface is created when the software used is just increasing the font weight, while a real boldface is a boldface designed for a specific

typeface. A real boldface is recognized by the title of the typeface, by adding the definition "bold" to the typeface name. In this experiment, all the selected typefaces have real boldface.

To avoid a possibly overwhelming task for the observers, it was necessary to limit the number of typefaces to ten. The selection criteria of typefaces were based on frequency of use and font family. The selected typefaces and their characteristics are listed in table 1.

Table 1. Typefaces included in experiment.

Typeface	Normal	Bold	Serif	Sans serif
Helvetica	X			X
Tiresias	X			X
Tiresias bold		X		X
Scala	X			
Scala bold		X	X	X
Scala sans	X		X	X
Scala sans bold		X		X
Verdana	X			
Times roman	X		X	
Frutiger	X			X

The text was printed on UPM Fine Woodfree paper with a gloss of 5 to 15, which is approximately comparable to the gloss found in printed newspapers. The contrasts are described as the percentage of black in the printed text and background. The contrasts in the experiment are listed in table 2.

Table 2. Contrasts in the experiment.

Text	Background
100 % black	0 % black
100 % black	20 % black
100 % black	40 % black
100 % black	60 % black
100 % black	80 % black
0 % black	20 % black
0 % black	40 % black
0 % black	60 % black
0 % black	80 % black
0 % black	100 % black

In the experiment, the font sizes are 8, 10, 12, 14 and 16. As mentioned earlier, the x height may vary for different typefaces in the same font size. For the typefaces and sizes to be comparable in the experiment, the x height of each of the typefaces was scaled using Times roman as a scaling reference. The x height of each typeface is scaled to match Times roman. This is illustrated in table 3, showing the font size, the x height and the scale factor for all typefaces.

Table 3. Typefaces and scaling factor based on Times roman.

Typeface	Scaling factor	8 pt	10 pt	12 pt	14 pt	16 pt
Helvetica	0,861	1,27/6,89	1,57/8,61	1,91/10,33	2,22/12,05	2,55/13,78
Tiresias	0,841	1,27/6,73	1,57/8,41	1,91/10,09	2,22/11,77	2,55/13,46
Tiresias bold	0,841	1,27/6,73	1,57/8,41	1,91/10,09	2,22/11,77	2,55/13,46
Scala	0,990	1,27/7,92	1,57/9,90	1,91/11,88	2,22/13,86	2,55/15,84
Scala bold	0,990	1,27/7,92	1,57/9,90	1,91/11,88	2,22/13,86	2,55/15,84

Scala sans	0,990	1,27/7,92	1,57/9,90	1,91/11,88	2,22/13,86	2,55/15,84
Scala sans bold	0,966	1,27/7,73	1,57/9,66	1,91/11,59	2,22/13,52	2,55/15,46
Verdana	0,826	1,27/6,61	1,57/8,26	1,91/9,91	2,22/11,56	2,55/13,22
Times roman	1,000	1,27/8	1,57/10,00	1,91/12	2,22/14	2,55/16
Frutiger	0,881	1,27/7,05	1,57/8,81	1,91/10,57	2,22/12,33	2,55/14,10

To summarize the parameters used in the experiment: 10 typefaces were selected, with characteristics as boldface/not boldface, serifs/non serifs, 10 variations of contrasts (including black type on white to dark grey background and white type on black to light gray backgrounds) and 5 font sizes. This gives a total of 500 variations.

3.2.2. Experiment method

The experiment was conducted as a home based experiment, where the booklets to be examined were distributed via postal service. The subjects received a booklet with the text and questionnaire. This booklet contained 50 pages with a text of minimum 3 text lines per sample with a line spacing factor of 120 %. A sample of the booklet is shown in figure 2.

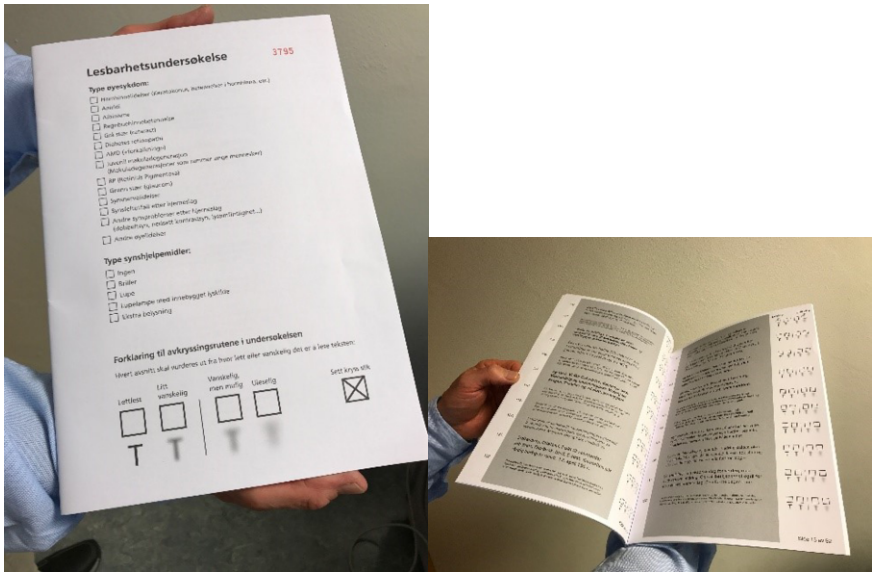

Figure 2. A booklet sample.

The number of sentences may vary for each paragraph. The observers were asked to read the text and scale the legibility based on four categories from “Easily readable”, “Readable with some difficulties”, “Difficult to read” to “Unreadable”. As this was a home based test, with the subjects in their home environment, aids like magnifying glass of supplementary lights were accepted.

4. Results and future work

Of the 5.000 subjects receiving the booklet and questionnaire, 830 responded. Data analysis show a systematic tendency in the data with little dispersion and many significant results. This is a strong proof that the chosen method is valid, otherwise the results would have been random. This was not the case, the results were indeed significant. This indicate that the noise factors of an uncontrolled experiment are compensated by a high number of test subjects.

Readability and legibility are affected by several parameters, more than the limited number of legibility parameters used in our experiment. Important factors include how the text itself is written, white space, leading, kerning, colors, paper quality and many more. Our focus was on the text as content, not headings or captions. For that reason, our study is concentrated on contrasts in the scale of black – white and simple line spacing. A focus on line spacing would increase the size of the booklet and lead to higher costs in printing the booklet, although it is established that line spacing does affect the readability.

The booklet contained 50 pages of text, as illustrated in figure 2. By including just one more parameter, the size of the booklet would increase to 100 pages. Participation in such a study require that the subjects do not withdraw because of boredom in the task or too large effort required. Regarding that the observer group included visually impaired subjects, reading is already a challenge. Having the volume of 830 subjects responding by assessing 500 experiment tasks is already remarkable. We did not want to try to push the limits by including more parameters, as we did not know the maximum limit of parameters for a subject to assess.

The subjects reported their own visual impairments. It is remarkable that the reported impairments contain such a variety of visual impairments. A possible explanation is that the variety of visual impairments is real, or that the question is not thoroughly understood. The list of visual impairments are long. A subject will most probably recognize her own visual impairments, and maybe other known visual impairments, but some expressions or names will not be recognized. Then this question could have been complicated and difficult to report in a home based uncontrolled experiment. For the results of the study, the only consequence of this is that the results are not able to be sorted based on visual impairments but on general basis. However, the goal of the study was to examine legibility in print for visually impaired subjects, and not related to specific impairments.

To conclude, the home based uncontrolled experiment method used in this study was a success, and the authors will recommend the method for similar studies.

The specific research results from the study is comprehensive, and will be published in international scientific journals.

5. Acknowledgements

This study was funded by the Delta Centre and Norges Blindeforbund. The experiment was conducted by the Norwegian Research Laboratory for Universal Design in collaboration with Norges Blindeforbund.

References

[1] J. Nersveen, *Forskningsprosjekt: Lesbarhet av trykt tekst for svaksynte og dyslektikere.* The Norwegian Research Laboratory for Universal Design. Gjøvik University College, 2013.
[2] O. Lund, *Knowledge construction in typography : the case of legibility research and the legibility of sans serif typefaces*, Doctoral dissertation thesis, The University of Reading, Department of Typography & Graphic Communication, 1999.
[3] Ø. Rannem, *Typografi og skrift*, Abstrakt forlag, 2012.
[4] G.E. Legge, C.A. Biggelow, Does Print Size Matter for Reading. A Review of Findings from Vision Science and Typography, *Journal of Vision* **11** (2011), 1-22.
[5] R. Bringhurst, *The Elements of Typographic Style*, Fourth edition, Hartley & Marks Publishers, 2012.
[6] L. Akhmadeeva, I. Tukhvatullin, B. Veytsman, Do serifs help in comprehension of printed text? An experiment with Cyrillic readers, *Vision Research* **65** (2012), 21-24.
[7] O. Lund, Why serifs (still) are important, *Typography Papers* **2** (1997), 91-104.
[8] M.A. Tinker, D. G. Paterson, Studies of typographical factors influencing speed of reading. VII. Variations in color of print and background, *Journal of Applied Psychology* **15** (1931), 471-479.
[9] N. Moroney, Unconstrained web-based color naming experiment, *Proc. SPIE vol. 5008, Color Imaging VIII: Processing, Hardcopy, and Applications*, Electronic Imaging Conference, Santa Clara, CA, January 2003.
[10] S. Zuffi, C. Brambilla, R. Eschbach, A. Rizzi, A study on the equivalence of controlled and uncontrolled visual experiments, *Proc. SPIE vol. 7241, Color Imaging XIV: Displaying, Processing, Hardcopy, and Applications,* Electronic Imaging Conference, pp. 724102.1-724102.11, San Jose, CA, January 2009.
[11] S. Zuffi, C. Brambilla, R. Eschbach, A. Rizzi, Controlled and Uncontrolled Viewing Conditions in the Evaluation of Prints, Proc. SPIE vol. 6807, *Color Imaging XIII: Processing, Hardcopy and Applications,* Electronic Imaging Conference, pp. 680714-680714-7, San Jose, CA, January 2008.