

Performance differences between Times and Helvetica in a reading task

RUDI W. DE LANGE, HENDRY L. ESTERHUIZEN AND DEREK BEATTY

*Department of Graphic Design
Technikon OFS
Private Bag X20539
Bloemfontein 9300, South Africa*

e-mail: rudi@studm.tofs.ac.za

SUMMARY

Typographers and printers often regard seriffed or roman typefaces as more legible and appropriate for reading material than typefaces without serifs. Authors contend that readers prefer roman above sans serif, that it is read faster, and that the comprehension rate is possibly higher when text is set in a roman typeface.

The absence of satisfactory empirical data to prove these assumptions, and the importance of legibility in academic reading material, motivated this study. The aim of the study was to determine the comparative legibility of sans serif and roman typefaces.

Four hundred and fifty primary school subjects from nine different schools were used in a control group pre-test, post-test research design where four different experiments were completed.

Romans and sans serifs were found to be equally legible, as no significant statistical difference was found between the reading speed, scanning speed, accuracy and comprehension at the 0.05 level.

These results are in contrast to the assumption that romans are more legible than sans serifs. They can be interpreted as promising for graphic designers and typographers, as it appears that legibility will not necessarily be sacrificed when certain reading material is set in a sans serif typeface.

KEY WORDS Legibility Sans serif typeface Roman typeface Reading task Times Roman Helvetica

1 INTRODUCTION

Typographers and printers traditionally regard typefaces with serifs as more legible than typefaces without serifs. Authors argue that readers prefer seriffed typefaces, read them faster, recognize them easier and that there could possibly be a higher comprehension rate with material printed in these typefaces. There is also the belief that seriffed typefaces assist in a horizontal movement whilst reading, and that the serifs help to distinguish different letters from each other [1, p. 30]. This alleged horizontal flow that is necessary for comfortable reading does not seem to concur with the saccadic eye movements that really occur during reading. Collier and Cotton [2, p. 34] believe that because serifs modulate

the space between letters, and because of the softer tone of romans, seriffed letters are more suitable for text. Rubinstein states that serifs provide ‘... *some visual spacing and accentuation to the ends of strokes that may help the reader read faster and avoid fatigue*’ [3, p. 16]. He also suggests that serifs make the ends of letters more visible, and the main strokes in the letters are thus easier to identify [3, p. 24]. Brady [4, p. 6] and Craig and Bevington [5, p. 89] are amongst some of the authors who believe that seriffed typefaces are more legible than typefaces without serifs. Most of these opinions are, however, not based on any satisfactory empirical evidence, but appear to reflect the personal opinion of the authors. The early roman letters designed for printing were based upon early handwritten manuscripts and the roman capitals carved on the buildings of the Roman Empire [6, p. 104]. It seems, therefore, logical to reason that traditional and historical factors played a major part in the custom of using typefaces with serifs for text in the Western world. The legibility theory of roman typefaces was developed after romans had been used a considerable time for text, and not because they are presumably superior in legibility. It was only in the beginning of this century that sans serifs were designed in greater numbers. According to McLean [7, p. 64] the first sans serif appeared in Britain in 1816 and the first lowercase sans serif in 1835.

Tinker [8, p. 46], reporting on earlier studies conducted by Paterson and himself, tested ten different typefaces for legibility using a speed reading test. Only one sans serif, Kable light, was included in this test. Kable light was read 2.2% slower than their standard type, Scotch Roman, and 2.6% slower than the fastest type, which was Garamond. Tinker regards this difference at the 0.05 level as ‘barely significant’ and states that ‘*A difference of about 3.6 per cent or greater tends to be significant at the 1 per cent level, i.e., highly significant*’ [8, p. 48].

In another interesting study a computer was used to simulate the neurological structure of the human visual system. A digital computer model of human visual processing was applied to characters with and without serifs. In the discussion the researchers concluded that serifs were important and useful in the perception of small individual characters [9, p. 359]). This experiment determined the importance of serifs in the perception of individual letters and not the legibility of continuous text.

There are also some authors who seem to suggest an equal legibility between roman typefaces and typefaces without serifs. In a related field of study, Poulton [10, p. 361] conducted an experiment to determine the comprehension rate of four sans serifs and three roman faces. No reliable difference was found between the comprehension rate of the roman and sans serifs. In his conclusion Poulton suggests: ‘... *it is not necessarily serifs, as has been claimed [Burt, 1959] which make typefaces readable.*’

Hartley [11, p. 8], in a paper discussing the role of print-based research in designing electronic text, mentions the following: ‘*Some investigators have argued that serifs in printed text increase the spacing between the letters slightly, and that this makes the text easier to read, but there does not seem to be any conclusive proof for such assumptions ...*’ White [12, pp. 12–17] concludes that it is impossible to make a clear-cut list of the advantages of a roman or sans serif typeface. He suggests that the reader must follow his own preference and that serifs are only one factor in the legibility of text.

Information provided by legibility research can guide the designer, printer and typographer to avoid factors that could diminish legibility. Designers and typographers can use results from legibility research to design and use typographical images to make them communicate more effectively in printed and electronic media.

2 THE PROBLEM

Assumptions about the superior legibility of seriffed or roman typefaces appear to be untested generalizations. The author found no supporting evidence during an extensive literature study to confirm them. Many typographical practices are still based on the belief that romans are the most legible typefaces to use for text. This unsubstantiated belief, the importance of legible instructional text, and the central part that typography plays in the graphic design process, motivated this study.

3 THE METHOD

The aim of this study was to determine the legibility of romans compared to typefaces without serifs. A word recognition test, speed reading test, comprehension test, and a scanning test were used to test the hypothesis that romans and sans serifs are equally legible. Accuracy, speed and comprehension were used as the criteria for legibility.

A systematic and random sample of four hundred and fifty primary school subjects from nine different primary schools was used in a word recognition, speed reading, comprehension marathon and a scanning test. The subjects were homogeneous regarding educational level and the language spoken at home. Only subjects with normal eyesight and those that did not receive academic remediation at the time of the study were included in these experiments.

A pre-test post-test control group research design was used for all the experiments. The level of significance was set at 0.05. The *t* test was used as the statistical test on the mean gain scores between the pre-tests and the post-tests to test the hypothesis that romans and sans serifs are equally legible. The alternative hypothesis stated that there is a difference between the two typefaces, as there was the possibility that sans serifs could be more legible.

4 THE RESULTS

The authors obtained sufficient evidence during the word recognition, speed reading, and comprehension test not to reject the research hypothesis of equal legibility between roman and sans serif typefaces. Serifs do not appear to have a noticeable effect on legibility, as measured by all the tests employed in this study. The subjects did not read and recognize words with serifs faster, their comprehension did not increase, and they were also not able to find a word in a portion of text more easily, when the text was set in a roman typeface. These tests and the corresponding results are fully discussed by de Lange [13]. Sufficient evidence was also found during the scanning experiment not to reject the hypothesis that Times Roman and Helvetica are equally legible. The author regards the scanning experiment as the most convincing experiment in this study. Scanning for a certain word or phrase does not require the reader to comprehend the text, that is, if the purpose of the scanning is to locate a particular section in the text, and not to grasp the meaning of the text. Competence in verbal fluency, or the lack of it, does not influence a reader's scanning ability, as this is a silent process. It seems to be reasonable to contend that the physical structure of the text and typeface would play the major part in a legibility experiment when scanning is used as a measure. If a researcher controls all the extraneous variables, then it is only the typographical attributes of the text that a reader must contend with whilst scanning. Details of the scanning experiment are as described in the following sections.

4.1 The scanning experiment

The authors systematically selected four primary schools, one from a higher socio-economic suburb, one from a lower, and two from middle-class suburbs.

Eighty scholars from Standard Two and eighty Standard Four scholars participated in this experiment. These are scholars in their fourth and sixth year of schooling, respectively.

All the scholars in the Standard Two and Standard Four classes were first systematically divided into two groups based upon their academic performance. Their class teachers determined their academic performance by the average achievement in their first language and their term average of the previous term. Scholars with scores in the upper half formed one group, and those in the lower half, the other group. Ten subjects were then randomly selected from the upper half and ten from the lower half from each standard and from each school. A set of random numbers, and an alphabetical class list provided by the schools, were used for this purpose. The subjects were then randomly divided into a control and experimental group. This provided forty subjects each in the control and experimental groups of the two standards. The author used this method in an attempt to match the experimental and control group more evenly, and possibly increase the experiment's sensitivity.

4.2 The material

Reading material for the Standard Two group was taken from Blanckenberg and Ferreira Blanckenberg [14, pp. 47–53], a discontinued school reading book for Standard Two. A discontinued school reader for Standard Four by the same authors provided the reading material for the Standard Four group [15, pp. 73–77]. Times Roman 13 point was used for the roman typeface and Helvetica 12 point for the sans serif typeface in this experiment. Type size, weight, line and letter spacing of the two faces were matched closely by means of computer software. The heights of the lower case letter x of Helvetica and Times Roman were respectively 2.30 mm and 2.25 mm. The height of the Times Roman capitals was 3.3 mm, and those of Helvetica 3.0 mm. Line spacing for both typefaces was 5.15 mm. The text was set flush left with a maximum line length of 33.75 picas. The pre-test consisted of 536 words and 47 lines and the post-test of 549 words and 45 lines.

4.3 The procedure

The subjects were informed that they were chosen to participate in an experiment with the aim of seeing how fast they could locate a particular word in a portion of text. The author provided the subjects with a full explanation of the scanning process. They did not have to comprehend the text, but were only required to find a particular word and mark it with a pencil.

The subjects first completed a trial scanning test to ensure that the correct procedures were followed and that they did not waste time in underlining or regressing after they had scanned the text. The experiments were conducted in empty classrooms or offices at the chosen schools.

Scanning times were recorded with a stopwatch. An index consisting of the percentage of correctly recognized words was multiplied with the scanning time and was taken as the score. The experimental group first received the pre-test set in a roman face (Times Roman), then the post-test set in a sans serif (Helvetica). The control group received their pre-test and post-test in Times Roman.

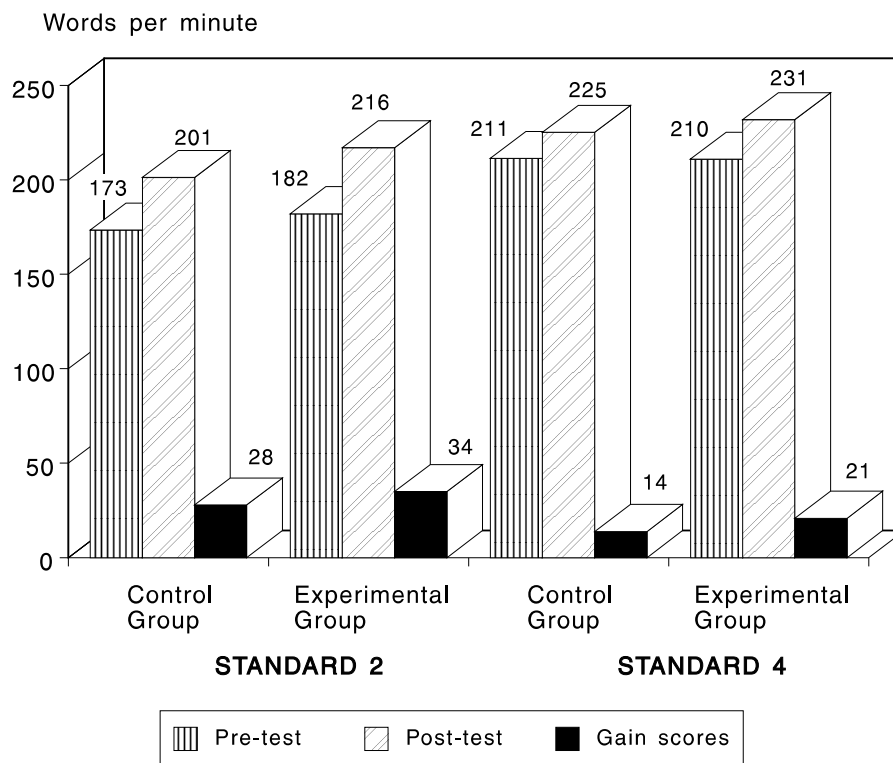


Figure 1. A graphical representation of the means and gain scores of the scanning test. (Times Roman compared to Helvetica)

4.4 The results of the scanning test

The results of the scanning test are illustrated graphically in Figure 1.

The Standard Two control group obtained a mean score of 173 words per minute in the pre-test and 201 in the post-test. The experimental group obtained a mean score of 182 and 216 in the pre-test and post-test respectively. This gave the control group a mean increase of 28 words per minute, whilst the experimental group's mean gain score increased to 34 words per minute.

With the Standard Four control group the mean increased from 211 for the pre-test to 225 words per minute for the post-test. The experimental group scored 210 and 231 in the two tests respectively. The control group's mean gain score was 14 words per minute and the experimental group's mean gain score 21 words per minute. These higher scores for the experimental groups are not statistically significant at the 0.05 level.

4.5 The computed t values

The computed t values are as follows: The Standard Two group: $t = -0.69$, the Standard Four group: $t = -0.567$. These values do not fall in the critical region, $t(0.05)2.00$. The null hypothesis is not rejected. There is no difference between the legibility of a sans serif and a roman typeface, when finding and counting a specific word in a scanning test is used as the criterion.

The lack of serifs for the experimental groups did not appear to influence the subjects' scanning abilities. Comprehension of the text was excluded, so it was only the typographical aspects of the text that a subject had to contend with. There was only one page in the pre-test and post-test, and this eliminated possible time differences between subjects when turning a page. Unknown factors which could influence the scanning indexes were the same for both groups, by nature of the research design.

It appears to be reasonable to conclude that there is no significant difference between the legibility of Times Roman and Helvetica as measured by a scanning process.

5 CONCLUSION AND DISCUSSION

The results of this study differ from the opinion of most authors on the subjects of typography, legibility and printing. These results can be interpreted as promising for designers and typographers, as it appears that legibility will not necessarily be sacrificed when sans serif typefaces are used for textual matter under the conditions of this study. Bluhm, writing on the development of a new typeface, Sassoon Primary Type and the related Sassoon Script Type and Sassoon Infant Type, reports that it has been developed to make reading and writing easier [16, p. 29]. This typeface is a sans serif and not a typeface with serifs. The results of this study appear to agree with the work of Zachrisson [17] and Moriarty and Scheiner [18, p. 703].

With normal primary school readers, and under normal reading conditions, sans serifs and romans can be regarded as equally legible. Serifs do not appear to affect legibility, as measured by the tests employed in this study. The authors are of the opinion that it is not necessarily serifs or the lack thereof that increase or decrease legibility. It is rather a complex interaction of known and unknown factors that affect a subject's reading performance and the legibility of reading material. Subject-matter, the readers' interest in the material, intellectual ability, and their emotional and physical condition, can all play a role in reading performance. The typefaces that the subjects are accustomed to appear to affect their typeface preference more than the typographical difference between romans and sans serifs.

Zachrisson [17, p. 115], in a series of experiments, used 48 children between ten and eleven years of age as a test group, and found no significant difference between romans and sans serifs in a silent speed reading comprehension test. He also stated: *'There is reason to assume that under normal conditions no significant difference exists between typefaces in common use by adults for running text'* [17, p. 36].

The results of the study reported in the present paper cannot be generalized to a wider population than that from which the samples were drawn. The authors are, however, of the opinion that readers older than the subjects will not find typefaces with serifs more legible than those without.

The unsubstantiated arguments that support the theory that romans are more legible than sans serifs can be opposed with satisfactory counter-arguments to support the authors' opinion. Some of these arguments and counter-arguments are as follows:

- Roman typefaces are more legible, because the theory states that serifs assist in the horizontal flow of reading and eye movements.

The counter-argument is that results from experiments using eye movement machines have shown that a reader's eyes do not flow but move in small steps. Serifs are, therefore, not required to assist in the reading process.

- Serifs increase spacing between letters and words and therefore aid perception and legibility.

The counter-argument is that it is possible to increase the letter space of any type style with modern typesetting equipment. Readers do not read by perceiving individual letters, but by recognizing individual words or portions of words.

- Serifs create a bigger irregularity in characters which helps to distinguish them from one another and are, therefore, more legible.

Readers, however, do not read individual characters, but rather fixate on words or portions of words when they read. Therefore, serifs are not necessary to create irregularities in characters.

- Serifs are strokes that bind characters into cohesive patterns and make it easier to recognize and read words set in a roman typeface.

Typefaces without serifs are normally set closer to each other, and serifs are not required to bind the characters into a cohesive pattern.

- Romans are predominantly used in reading material, and it is because of the familiarity aspect that seriffed faces are more legible. Both Turnbull and Baird [19, p. 86] and McLean [7, p. 44] mention familiarity as well as the irregular design features of serif faces as reasons for improved legibility.

The above argument is possibly one with some validity. There is, however, the question whether readers are consciously aware of the typographical difference between roman and sans serif when reading a book or newspaper. Ernst believes that serifs serve no useful purpose in the discernment of type, and that what a person is accustomed to plays a bigger role [20, p. 40].

- The old-style roman faces are less uniform in their proportions than some sans serifs, and it is argued that these uneven proportions aid legibility.

Some sans serifs are also based on old-style proportions, for example Optima. If uneven proportions increase or aid legibility, then old-style sans serifs could be just as legible as old-style romans.

- A strong argument that seems to support the theory of the superior legibility of romans is that the upper parts of words set in a roman typeface are easier to recognize than when they are set in a sans serif typeface.

Text is not usually presented with the lower half of the letters obscured, and words in text are also read in context with other words. An argument that seems to be convincing with one word cannot be generalized to text.

- The lack of serifs is said to contribute to a vertical stress in sans serifs. This vertical stress is supposed to compete with the horizontal flow of reading. It is therefore argued that romans are the more legible type, and that increased line spacing is required when sans are used for text type [21, p. 129].

Reading is, however, not a flowing movement. As early as 1878, Professor Javal of the University of Paris established that a reader's eyes move along a line of print in a series of quick jerks [22, p. 13]. This was contrary to the popular belief that a reader's eyes move along a line of print in a smooth sweep. Javal called these quick jerks saccadic movements. Rayner and Pollatsek [23, pp. 113–123] provide a comprehensive discussion on the work of the eyes during the reading process.

Other typographical and uncontrollable factors, for example, line spacing, the interest shown by the reader, and the quality of print or visual display units, play a bigger role in legibility than the typographical differences between sans serifs and roman typefaces.

The need for serifs in typeface design is questioned if its main purpose is to make typefaces more legible.

Could it be that ‘... serifs are the untidy mistakes of poor workmanship petrified into a traditional form!’ [12, p. 12].

ACKNOWLEDGEMENTS

The authors would like to acknowledge Technikon OFS for providing financial assistance to complete the experimental work, and the Orange Free State Education Department for granting permission to use scholars from primary schools under their control for the experiments.

REFERENCES

1. M. Beaumont, *Type & Colour*, Phaidon Press Limited, Oxford, 1987.
2. D. Collier and B. Cotton, *Basic Desktop Design and Layout*, Quarto Publishing, London, 1989.
3. R. Rubinstein, *Digital Typography: an Introduction to Type and Composition for Computer System Design*, Addison-Wesley, Reading, Massachusetts, 1988.
4. P. Brady, *Using Type Right*, North Light Books, Cincinnati, 1988.
5. J. Craig and W. Bevington, *Working with Graphic Designers*, Watson-Guptill Publishers, New York, 1989.
6. P. B. Meggs, *A History of Graphic Design*, Penguin Books, New York, 1983.
7. R. McLean, *The Thames and Hudson Manual of Typography*, Thames and Hudson, London, 1980.
8. M. A. Tinker, *Legibility of Print*, Iowa State University Press, Ames, Iowa, 1963. 3rd edition.
9. D. O. Robinson, M. Abbamonte, and S. H. Evans, ‘Why serifs are important: the perception of small print’, *Visible Language*, **4**, 353–359, (1971).
10. E. C. Poulton, ‘Letter differentiation and rate of comprehension in reading’, *Journal of Applied Psychology*, **49**(5), 358–362, (1965).
11. J. Hartley, ‘Designing electronic text: the role of print-based research’, *Educational Communication and Technology*, **35**(1), 3–17, (1987).
12. J. V. White, *Graphic Design for the Electronic Age*, Watson-Guptill Publishers, New York, 1988.
13. R. W. De Lange, *The Legibility of Sans Serif Typefaces, an Experimental and Comparative Study*, M. dip. technology dissertation, Technikon OFS, Bloemfontein, 1993.
14. N. A. Blanckenberg and A. M. Ferreira Blanckenberg, *Land, Lug en See. Moedertaal-leesboeke vir die Laerskool. St. 2*, Maskew Miller, Kaapstad, (n.d.).
15. N. A. Blanckenberg and A. M. Ferreira Blanckenberg, *Klinkende Dade. Moedertaal-leesboeke vir die Laerskool. St. 4*, Maskew Miller, Kaapstad, (n.d.).
16. A. Bluhm, ‘Types for easy reading’, *Graphic Repro*, **11**(3), 29–31, (1991).
17. B. Zachrisson, *Studies in the Legibility of Printed Text*, Almqvist & Wiksell, Stockholm., 1965.
18. S. E. Moriarty and E. C. Scheiner, ‘A study of close-set text type’, *Journal of Applied Psychology*, **69**(4), 700–703, (1984).
19. A. T. Turnbull and R. N. Baird, *The Graphics of Communication*, Holt, Rinehart and Winston, New York, 1980. Fourth edition.
20. S. B. Ernst, *The ABC's of Typography*, Art Direction Book Company.
21. J. Craig, *Designing with Type*, Watson-Guptill Publishers, New York, 1981.
22. H. Spencer, *The Visible Word*, Lund Humphries in association with the Royal College of Art, London, 1969.
23. K. Rayner and A. Pollatsek, *The Psychology of Reading*, Prentice-Hall Inc., Englewood Cliffs, 1989.