Typeface Legibility: 
Towards defining familiarity

Sofie Beier

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

May 2009

The Royal College of Art
© This text represents the submission for the degree of Doctor of Philosophy at the Royal College of Art. This copy has been supplied for the purpose of research for private study, on the understanding that it is copyright material, and that no quotation from the thesis may be published without proper acknowledgement.
Abstract

The aim of the project is to investigate the influence of familiarity on reading. Three new fonts were created in order to examine the familiarity of fonts that readers could not have seen before. Each of the new fonts contains lowercase letters with familiar and unfamiliar skeleton variations. The different skeleton variations were tested with distance threshold and time threshold methods in order to account for differences in visibility. This investigation helped create final typeface designs where the familiar and unfamiliar skeleton variations have roughly similar and good performance. The typefaces were later applied as the test material in the familiarity investigation.

Some typographers have proposed that familiarity means the amount of time that a reader has been exposed to a typeface design, while other typographers have proposed that familiarity is the commonalities in letterforms. These two hypotheses were tested by measuring the reading speed and preference of participants, as they read fonts that had either common or uncommon letterforms, the fonts were then re-measured after an exposure period. The results indicate that exposure has an immediate effect on the speed of reading, but that unfamiliar letter features only have an effect of preference and not on reading speed.

By combining the craftsmen’s knowledge of designing with the methods of experimental research, the project takes a new step forward towards a better understanding of how different typefaces can influence the reading process.
Contents

- Copyright statement ........................................ 002
- Abstract ..................................................... 003
- Contents .................................................... 004
- List of illustrations .......................................... 007
- Acknowledgements ......................................... 009
- Author’s Declaration ....................................... 010

1. Setting the Stage
   1.1. Introduction ........................................... 012
   1.2. Different approaches to legibility ................. 015
       - The two cultures
       - The designer’s lack of interest
   1.3. Defining the meaning of central terms and use of language .......................... 021
   1.4. The validation of the most essential test methods ....................... 024
       - Continuous reading
       - Reading aloud
       - Errors
       - Speed
       - The search task
       - Visual accuracy threshold
       - Variable distance
       - Short exposure
       - Readers’ preferences
   1.5. Perception of letters ................................ 034
       - Letter identification
       - Single letter and word superiority effects
       - Word wholes
       - Parts, wholes & context
   1.6. The Internal relation of visibility and familiarity .................................... 044
       - The cognitive aspect of separation in visual perception
       - ‘Readers read best what they read most’
2. Visibility

2.1. The visibility of typefaces for different needs
- Continuous reading
- Type at a distance
- The pixel
- On print

2.2. Empirical findings on differentiation of lowercase characters
- Individuality of the x-height letters
- Individuality of the narrow letters

2.3. Designers and differentiation of lowercase characters
- Differentiation in the early days of printing
- Individuality of the x-height letters
- Individuality of the narrow letters

2.4. Three new fonts

2.5. A Study of the character shapes in three new fonts
- Test material
  - Familiar letter variations
  - Unfamiliar letter variation
- Study 1
  - Methods
    - Participants
    - Material
    - Procedures
  - Results of short exposure study
- Study 2
  - Methods
    - Participants
    - Material
    - Procedures
  - Results of distance study
- Discussion
  - The x-height letters
  - The narrow letters
  - The unfamiliar letter shapes

2.6. Implementing results in three new typefaces
3. Familiarity

3.1. The long-term influence of familiarity 110
   • Uncommon letterforms
     • Reading and writing in the Middle Ages
     • Gothic and Latin types
   • Common letterforms
     • John Baskerville
     • Didone
     • Serif and Sans Serif

3.2. A study of familiarity before and after an exposure session 123
   • Test material
   • Method
     • Participants
     • Material
     • Procedures
   • Results
     • Reading speed for group-1
     • Reading speed for group-2
     • Questionnaires
   • Discussion
     • The exposure hypothesis
     • The letter feature hypothesis

4. Closure

4.1 Summary, perspectives & conclusion 142
   • Future familiarity studies
   • Future visibility studies

Appendix A: Data of visibility study 152
Appendix B: Tinker sketches & specimen 160
Appendix C: Ovink sketches & specimen 176
Appendix D: Pyke sketches & specimen 194

Definitions & Bibliography

• Letter-Parts 210
• Glossary 211
• Bibliography 218
Figures

<table>
<thead>
<tr>
<th>no.</th>
<th>Figure Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Stroop effect</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>The internal relationship of letters in the feature-comparison theory</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>Neisser’s lists for visual searching (Neisser 1967)</td>
<td>36</td>
</tr>
<tr>
<td>4</td>
<td>Explaining the forced-choice study</td>
<td>37</td>
</tr>
<tr>
<td>5</td>
<td>The Parallel Letter Recognition (PLR) model</td>
<td>41</td>
</tr>
<tr>
<td>6</td>
<td>The knockout test material by Pelli &amp; Tillman (2007)</td>
<td>42</td>
</tr>
<tr>
<td>7</td>
<td>Visualisation of test material by Hochberg (1968)</td>
<td>44</td>
</tr>
<tr>
<td>8</td>
<td>Data from the study by Posner &amp; Mitchell (1967)</td>
<td>46</td>
</tr>
<tr>
<td>9</td>
<td>The typefaces InfoText and InfoDisplay by Ole Schäfer and Erik Spiekermann</td>
<td>51</td>
</tr>
<tr>
<td>10</td>
<td>The typeface Balance</td>
<td>52</td>
</tr>
<tr>
<td>11</td>
<td>The typeface Quadraat</td>
<td>52</td>
</tr>
<tr>
<td>12</td>
<td>The typeface Trinité</td>
<td>52</td>
</tr>
<tr>
<td>13</td>
<td>The typeface Avance</td>
<td>52</td>
</tr>
<tr>
<td>14</td>
<td>Dwiggins’ illustration of the M-formula doll</td>
<td>53</td>
</tr>
<tr>
<td>15</td>
<td>Dwiggins’ m-formula in type design</td>
<td>53</td>
</tr>
<tr>
<td>16</td>
<td>The fonts tested by Waller (2007)</td>
<td>54</td>
</tr>
<tr>
<td>17</td>
<td>The ‘Series B’ and the ‘Series D’, the former standard fonts for American Highways</td>
<td>54</td>
</tr>
<tr>
<td>18</td>
<td>Sketches by Gerard Unger for M.O.L.</td>
<td>55</td>
</tr>
<tr>
<td>19</td>
<td>The typefaces Akzidenz and Airport</td>
<td>56</td>
</tr>
<tr>
<td>20</td>
<td>The typeface Univers</td>
<td>56</td>
</tr>
<tr>
<td>21</td>
<td>Signage from the Charles de Gaulle Airport</td>
<td>56</td>
</tr>
<tr>
<td>22</td>
<td>The typeface Cheltenham</td>
<td>57</td>
</tr>
<tr>
<td>23</td>
<td>Test material applied Roethlein (1912)</td>
<td>57</td>
</tr>
<tr>
<td>24</td>
<td>Example of hinting in the typeface Georgia</td>
<td>58</td>
</tr>
<tr>
<td>25</td>
<td>Cleartype in the typeface Constantia</td>
<td>58</td>
</tr>
<tr>
<td>26</td>
<td>Optical scaling by Justus Erich Walbaum</td>
<td>58</td>
</tr>
<tr>
<td>27</td>
<td>Type by Johann Fleischman</td>
<td>59</td>
</tr>
<tr>
<td>28</td>
<td>The typeface Parable</td>
<td>59</td>
</tr>
<tr>
<td>29</td>
<td>The typeface Bell Centennial</td>
<td>59</td>
</tr>
<tr>
<td>30</td>
<td>The typeface Optima</td>
<td>60</td>
</tr>
<tr>
<td>31</td>
<td>The typefaces Meta and Officina</td>
<td>60</td>
</tr>
<tr>
<td>32</td>
<td>The typeface Gill Sans</td>
<td>60</td>
</tr>
<tr>
<td>33</td>
<td>The version of Futura applied by Geyer (1977)</td>
<td>62</td>
</tr>
<tr>
<td>34</td>
<td>The work of five different researchers looking into misreading of lowercase letters</td>
<td>63</td>
</tr>
<tr>
<td>35</td>
<td>Test material applied by Legros &amp; Grant (1916)</td>
<td>64</td>
</tr>
<tr>
<td>36</td>
<td>The test material applied by Ovink (1938)</td>
<td>65</td>
</tr>
<tr>
<td>37</td>
<td>Letter ‘a’ variations of the study carried out by Spencer and colleagues (1973)</td>
<td>65</td>
</tr>
<tr>
<td>38</td>
<td>The test material applied by Harris (1973)</td>
<td>65</td>
</tr>
<tr>
<td>39</td>
<td>The typeface Didot</td>
<td>66</td>
</tr>
<tr>
<td>40</td>
<td>The typeface Adobe Jenson</td>
<td>66</td>
</tr>
<tr>
<td>41</td>
<td>The typeface Adobe Garamond</td>
<td>66</td>
</tr>
<tr>
<td>42</td>
<td>The typeface Swift</td>
<td>67</td>
</tr>
<tr>
<td>43</td>
<td>The typeface Didot</td>
<td>67</td>
</tr>
<tr>
<td>44</td>
<td>Type specimen of work by Van Dijck</td>
<td>70</td>
</tr>
<tr>
<td>45</td>
<td>Type specimen of work by Fournier</td>
<td>70</td>
</tr>
<tr>
<td>46</td>
<td>Type specimen of work by Caslon</td>
<td>71</td>
</tr>
<tr>
<td>47</td>
<td>Characters from a Bodoni type specimen</td>
<td>72</td>
</tr>
<tr>
<td>48</td>
<td>Type specimen of work by Garamond</td>
<td>72</td>
</tr>
<tr>
<td>49</td>
<td>Type specimen of work by Didot</td>
<td>72</td>
</tr>
<tr>
<td>50</td>
<td>Type specimen of work by Griffo</td>
<td>73</td>
</tr>
<tr>
<td>51</td>
<td>Type by Jenson</td>
<td>73</td>
</tr>
<tr>
<td>52</td>
<td>Type specimen of work by Goudy</td>
<td>73</td>
</tr>
<tr>
<td>53</td>
<td>The typeface Clearface</td>
<td>73</td>
</tr>
<tr>
<td>54</td>
<td>The typeface Futura</td>
<td>74</td>
</tr>
<tr>
<td>55</td>
<td>The typeface OCR B</td>
<td>74</td>
</tr>
<tr>
<td>56</td>
<td>The London Transport typeface</td>
<td>74</td>
</tr>
<tr>
<td>57</td>
<td>Johnston’s ‘the Foundation Hand’</td>
<td>74</td>
</tr>
<tr>
<td>no.</td>
<td>page</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>A Romain du Roi plate .................................................. 75</td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>Psalter printed by Fust and Schoeffer .................................. 75</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>A calligrapher specimen possibly Nicolas Jarry ......................... 75</td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>The typefaces Centaur and Adobe Jenson .................................. 75</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>Type specimen of work by Caslon ........................................... 75</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>The terminal of the lowercase ‘f’ on the type body ....................... 76</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>The terminal of the lowercase ‘f’ ........................................ 76</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>The typeface Sabon .......................................................... 76</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>The typeface Sabon Next .................................................... 76</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>Type specimen of work by Garamond ...................................... 77</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>Average number of paragraphs read for group-2 ........................... 132</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>Average number of paragraphs read for group-1 ........................... 131</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>Test material for group-2 .................................................. 127</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>Different levels of typeface familiarity .................................. 126</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>A common basic form of typefaces by Adrian Frutiger ....................... 124</td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>Type specimen of work by Bodoni ......................................... 81</td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>Skeleton variations for test material ..................................... 84</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>The TinkerTest letter variations ........................................... 85</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>The OvinkTest letter variations ............................................ 85</td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>The PykeTest letter variations .............................................. 85</td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>Visualisation of the exposure method ...................................... 87</td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>Errors in study 1 for the letter ‘a’ ....................................... 88</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>Errors in study 1 for the letter ‘c’ ....................................... 88</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>Errors in study 1 for the letter ‘e’ ....................................... 89</td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>Errors in study 1 for the letter ‘n’ ....................................... 89</td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>Errors in study 1 for the letter ‘s’ ....................................... 89</td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>Errors in study 1 for the letter ‘u’ ....................................... 90</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>Errors in study 1 for the letter ‘f’ ....................................... 90</td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>Errors in study 1 for the letter ‘i’ ....................................... 90</td>
<td></td>
</tr>
<tr>
<td>87</td>
<td>Errors in study 1 for the letter ‘j’ ....................................... 91</td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>Errors in study 1 for the letter ‘l’ ....................................... 91</td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>Errors in study 1 for the letter ‘t’ ....................................... 91</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>Errors in study 1 for the letter ‘a’ ....................................... 93</td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>Errors in study 2 for the letter ‘c’ ....................................... 93</td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>Errors in study 2 for the letter ‘e’ ....................................... 94</td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>Errors in study 2 for the letter ‘n’ ....................................... 94</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>Errors in study 2 for the letter ‘s’ ....................................... 95</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>Errors in study 2 for the letter ‘u’ ....................................... 95</td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>Errors in study 2 for the letter ‘f’ ....................................... 95</td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>Errors in study 2 for the letter ‘i’ ....................................... 96</td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>Errors in study 2 for the letter ‘j’ ....................................... 96</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>Errors in study 2 for the letter ‘l’ ....................................... 96</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>Errors in study 2 for the letter ‘t’ ....................................... 97</td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>The lowercase ‘f’ in TinkerRegularRoman ................................... 103</td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>The typefaces PykeMicro, PykeText and PykeDisplay ......................... 103</td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>Overview of all fonts in the Tinker typeface family ......................... 105</td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>Overview of all fonts in the Ovink typeface family ......................... 106</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>Overview of all fonts in the Pyke typeface family .......................... 107</td>
<td></td>
</tr>
<tr>
<td>106</td>
<td>Paul Renner’s first designs for Futura .................................... 111</td>
<td></td>
</tr>
<tr>
<td>107</td>
<td>Specimen of Half-Uncial Roman .............................................. 112</td>
<td></td>
</tr>
<tr>
<td>108</td>
<td>English Carolingian .......................................................... 113</td>
<td></td>
</tr>
<tr>
<td>109</td>
<td>Detail from the Mrs. Eaves type specimen ................................... 116</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>Type specimen of work by Didot ............................................. 118</td>
<td></td>
</tr>
<tr>
<td>111</td>
<td>Type specimen of work by Bodoni .......................................... 119</td>
<td></td>
</tr>
<tr>
<td>112</td>
<td>Test material applied by Miles A. Tinker ................................... 120</td>
<td></td>
</tr>
<tr>
<td>113</td>
<td>A common basic form of typefaces by Adrian Frutiger ....................... 124</td>
<td></td>
</tr>
<tr>
<td>114</td>
<td>Different levels of typeface familiarity ................................... 126</td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>Test material for group-1 .................................................... 127</td>
<td></td>
</tr>
<tr>
<td>116</td>
<td>Test material for group-2 .................................................... 127</td>
<td></td>
</tr>
<tr>
<td>117</td>
<td>Average number of paragraphs read for group-1 ........................... 132</td>
<td></td>
</tr>
<tr>
<td>118</td>
<td>Average number of paragraphs read for group-2 ........................... 132</td>
<td></td>
</tr>
<tr>
<td>119</td>
<td>The average response to the statement Q1 .................................. 133</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>The average response to the statement Q2 .................................. 134</td>
<td></td>
</tr>
<tr>
<td>121</td>
<td>The average response to the statement Q3 .................................. 135</td>
<td></td>
</tr>
<tr>
<td>122</td>
<td>The average response to the statement Q4 .................................. 135</td>
<td></td>
</tr>
<tr>
<td>123</td>
<td>The average response to the statement Q5.1 ................................ 136</td>
<td></td>
</tr>
<tr>
<td>124</td>
<td>The average response to the statement Q5.2 ................................ 136</td>
<td></td>
</tr>
</tbody>
</table>
Acknowledgements

I AM INDEBTED TO my external supervisor Kevin Larson, for essential guidance, statistical calculations and substantial help on the methodology in the creation of chapters 2.5 and 3.2.

My thanks go to James Mosley for useful comments on drafts of chapters 2.3 and 3.1, and for making me aware of the spur on the lowercase ‘l’ in Textura Gothic and French calligraphy.

I would like to give a special thanks to Gerard Unger for stimulating typographical conversations and comments on the final draft. I would also like to thank Linda Reynolds and Gerry Leonidas for pointing me in the right direction at the early stages of the project. My thanks further go to Per Galle, and the rest of the research group at The Danish Design School, for useful comments in the later stages. I must also thank Henrik Kubel for being highly critical and questioning my work, which resulted in me working even harder.

A special thanks goes to Harriet Edwards for helping me understand the nuances of the English language, and to Philip Swift for proofreading the final draft.

I would like to thank Ole Lund, Scott Bondurant Chandler and Rob Hillier for supplying me with copies of their PhD thesis, Karin von Ompteda for sharing references, and the staff at the British Library, the National Art Library, and St. Bride Library for providing the most incredible material.

Furthermore I thank the Microsoft Advanced Reading Technologies and the Rootstein Hopkins Foundation for financial support to the project. Last but not least, I am grateful to the Royal College of Art, my internal supervisor Jeff Willis, and the rest of the staff at the department of Communication Art and Design for making this research possible.
During the period of registered study in which this thesis was prepared the author has not been registered for any other academic award or qualification.

The material included in this thesis has not been submitted wholly or in part for any academic award or qualification other than that for which it is now submitted.
Section 1

SETTING THE STAGE
1.1

Introduction

The investigation of what constitutes legibility in type presents many difficulties, for it depends essentially on two main heads: (1) the subjective, or the view of the reader; and (2) the objective, or the features presented by the printed surface to be read (Legros, 1922, p.1).

Lucien Alphonse Legros, one of the authors of the highly regarded Typographical Printing Surfaces, suggested this in the introduction to a note for a committee appointed to select the best typefaces for government printing in England.

Typographers often talk about the importance of familiarity, yet familiarity has multiple meanings. For some, the term refers to the amount of time readers spend reading a particular typeface. For others, the term refers to the similarity between the features of a typeface and the features of other typefaces. It appears that typeface familiarity consists both of issues of exposure and of common features, but exactly how it does so is not yet properly understood.

While controlling the factors of the visibility of the type, the focus of the present project will be on the influence typeface familiarity has on the reader both on objective and subjective
levels. The aim is to investigate the question of whether a familiarity effect can be measured through an exposure time of reading typefaces of different levels of common letter features and of previous exposure. For the test material, three fonts were newly created that readers would never have seen before. Prior to the familiarity investigation early versions of the fonts were tested to control visibility. We know that visibility and familiarity are two separate operations of the perceptual system, and that visibility is an important characteristic for reading performance, so this parameter should therefore be controlled in order to study familiarity. Yet, to fully control the visibility of type, issues such as character differentiation, spacing, contrast, stroke, weight, width, resolution, hinting, and kerning, should ideally be dealt with simultaneously. However due to the limitation of space in the present project, the main focus of the visibility section is on the differentiation of characters. The other related topics are to be dealt with thoroughly in future research.

After setting the stage in Section 1, the structure of the project goes as follows: Section 2 begins with a presentation of selected material on typeface visibility, including an overview of the different usage of typefaces and a presentation of existing design recommendations related to each area. This is followed by reviews of scientific findings and of type designers’ approaches toward letter differentiation. Next is a presentation of the design of preliminary versions of three new fonts, all including several familiar and unfamiliar skeleton variations of the most frequently misrecognized lowercase letters. After this, there is a presentation of the new experimental investigation concerning visibility of single letters with the new fonts applied as test material. A distance threshold method and a time threshold method were used to study the visibility of alternate forms of different letters within each of the three fonts. The Section finishes with the implementation of the findings of the visibility investigation in the final versions of the three typeface families. Section 3 begins with an overview of historical references to typeface familiarity. The new experimental investigation concerning typeface familiarity will be presented. Three kinds of typefaces were compared against each other in each of two studies: an already familiar typeface, a new typeface with common
letter features, and a new typeface with uncommon letter features. Reading speed performance and opinions were measured both before and after an exposure period. The thesis finishes in Section 4 with a conclusion in which implications for future research are discussed.

---

In a review of work presented by legibility researchers and type designers, the project intends to unite skills from the two communities by placing practice on an equal footing with findings reached by scientific methods. In doing so, the ambition is to generate new knowledge that not only enlightens type designers in their work, but also educates academics working within the field of vision research. While type designers will be presented with legibility matters directly related to the craft, and introduced to different test methods and views on reading, the work will also inform the cognitive science community on the different influences typographical based test material can have on experimental studies.

By investigating how visibility functions in relation to the differentiation of the most often misread lowercase characters of three new fonts, and how familiarity functions in relation to known and unknown typefaces, the objective at a general level is to contribute to the field of legibility research, and on a practical level to provide tools usable in the design of new typefaces. By combining scientific and practice-based approaches – in reviews of empirical studies and in presentations of ideas and opinions put forward by designers – the project will give an overview of what we know about the legibility of the lowercase Latin alphabet. With that as a foundation, I will add to the existing body of knowledge by focusing on typeface familiarity in new experimental investigations.
1.2

Different approaches to legibility

A significant amount of experimental legibility research was carried out in the first half of the 20th century. From the late sixties, attention to these issues moved from the aim of improving reading matters towards an interest in cognitive processes of reading. In recent times, academic interest in legibility has been typically concerned with the new media, and often with a more cognitive approach than the one applied by early predecessors. A large portion of both past and present research on legibility is still relevant for typeface- and graphic designers. However, it is not an uncommon view among designers to see legibility issues as a limitation to their creative work, and to claim that the readers eventually will become used to whatever typeface they are presented with. Furthermore, the lack of easily accessible references often leads designers to make up their own assumptions based on intuition rather than knowledge.

The two cultures

In 1959 the scientist and novelist C.P. Snow published the influential book The Two Cultures. Here he described his own experience of a lack of communication between the sciences and the humanities. This polarisation, he argued, originates in a
mutual lack of understanding of how the other field contributes to knowledge:

The non-scientists have a rooted impression that the scientists are shallowly optimistic, unaware of man’s condition. On the other hand, the scientists believe that the literary intellectuals are totally lacking in foresight, peculiarly unconcerned with their brother men, in a deep sense anti-intellectual, anxious to restrict both art and thought to the existential moment (Snow, 1959, p.5).

This lack of understanding of other trades appears also to be present in relation to knowledge on legibility – an area dominated by the two cultures of the empirical scientific approach and of the more design-orientated approach. The empirical approach is based on the testing of hypotheses. It is a problem-solving approach where assumptions are tested with an outcome of some form of conclusion. The experiments are often carried out by people with a psychological and scientific background, but without any real knowledge of the design aspect of typography. According to the Oxford English Dictionary, the word empiricism is defined as: ‘the theory that all knowledge is derived from experience and observation’ (Oxford Dictionary, 2008, p.326), in contrast to knowledge based solely on thinking and reasoning. A scientific approach is based on examination, and only theories that are testable by empirical methods will be of interest. As put by Keith E. Stanovich:

Science is a mechanism for continually challenging previously held beliefs by subjecting them to empirical tests in such a way that they can be shown to be wrong (Stanovich, 2007, p.26).

A design-orientated approach, on the other hand, is more informal and is obtained by craftsmen through practice. Many designers have communicated the trade concepts with great success. Concepts and theories can be articulated verbally; however, the undefined matter of experience is a far more difficult issue to communicate. This phenomenon – described as tacit knowing – has been of interest to the scientist and philosopher Michael
Polanyi, who gives this example of riding a bike:

I cannot say clearly how I ride a bicycle [...] I know that I know perfectly well how to do such things, though I know the particulars of what I know only in an instrumental manner and am focally quite ignorant of them (Polanyi, 1974, p.88)

The argument is that we know much more than we are capable of communicating. According to Allan Janik (1988), this exemplifies the limits of science, his point being that any issue known tacitly alone cannot be studied scientifically. Nonetheless, he argues that the fact that we are incapable of communicating something does not imply that it is unknowable – as long as we can identify a certain type of knowledge, it must be known.

The designer’s lack of interest

It is possible that this difficulty in communicating tacit knowledge is partly responsible for many craftsmen and scientists seeming not to benefit from each other in sharing legibility related information. Scientists have sometimes carried out tests on a weak typographic basis1 dictated by the manufactures of type, this has led type designers to ignore the results and, as a consequence, to find a justification for carrying on as usual. This frustrated the productive legibility researchers Tinker2 and his associate Paterson:

Unfortunately, insight concerning one’s own ignorance of the complexity of the problem is rarely encountered in

---

1) One of the most common typographical weaknesses has to do with the optical size of a typeface that varies dramatically according to the size of the x-height character — generally understood among most designers, and a common topic in the majority of the literature introducing typography. However, the issue has either not been a known fact to some academic researchers, or they have chosen to ignore its importance due to the technical limitations in the reproduction of the applied typefaces; the problem has later been pointed out by several researchers (Zachrisson, 1965; Wiggins, 1967; Poulton, 1972; Lund, 2002). To make matters even more complicated, it is not uncommon that the type size on the body varies significantly between typefaces – a variation that has existed since the early printers (Burnhill, 2003). Roethlein (1932) and Tinker (1944) both highlighted the problem with regards to hand compositing. Some years later, and probably in reference to phototypesetting, Zachrisson (1965) pointed towards a difference in body size being as much as 25%, and Hermann Zapf (1987) gave examples of an uppercase difference of 40%.

2) The psychologist Miles Albert Tinker, in collaboration with colleagues Donald G. Paterson and Helen A. Webster, tested, over the course of his career, more than 30,000 subjects in various areas related to legibility and readability of typography.
the practical printer or the type designer. For this reason, practical printers and type designers are likely to disregard findings from the research laboratory or at best to accept only those findings which happen to coincide with their own beliefs (Tinker & Paterson, 1949, p.61).

In earlier years, a great amount of well-documented legibility experiments were carried out, not only by Tinker and his collaborators, but also by people such as Roethlin (1912), Pyke (1926), Kerr, Crosland & Johnson (1928), Luckiesh & Moss (1938), and the typographic scholar Gerrit Willem Ovink (1938).

The rejection by the practice world of this prior research also disturbed the former student of Edward Johnston, John C. Tarr:

The work already done by the scientist in the field of visibility and readability nevertheless deserves closer recognition from the printing industry. It is submitted that such interest would be wider if the scientist were to conduct his investigations on those practical lines that the printing industry would be only too willing to suggest (Tarr, 1949, p.31).

The lack of collaboration is still present in more contemporary design work. In 1999, Rick Poynor commented on the long list of earlier publications and studies on eye movements, reading speed and letter recognition:

It's a safe bet that few graphic designers will have read, or even heard of, many of these studies, or have any knowledge of the research material published since. To designers, the scientific approach seems fundamentally hostile to the mysteries of the creative process (Poynor, 1999, p.14).

Attempting to explain this lack of interest among designers, Merald Wrolstad, the editor of the Journal of Typographic Research (1969), emphasised that since the focus of any research project will always be motivated by the researcher's area of interest, and because the majority of legibility researchers are either psychologists, engineers or reading specialists, the research conducted will be concerned with these topics and not with letterforms. The findings will not be published in design journals,
but in journals related to the field of the researcher in question and in a format unfamiliar to the majority of designers. A different approach when analyzing results also seems to influence designers’ interest in the material. As an editor, Wrolstad sometimes struggled to convince design writers that in research ‘what did not work may sometimes be as important as what did work’ (Wrolstad, 1969, p.118).

Although recognizing the importance of legibility research with regards to specific matters such as creating typefaces for low resolution print-outs and helping children to read, the renowned typographer and writer Walter Tracy expressed his scepticism by stating that, in terms of the general matter of typography, most designers ‘know that the common element in it all is the familiar alphabet’ (Tracy, 1988, p.84). As a result, Tracy argued that typographers are mostly indifferent to the existing legibility research.

Not all academic researchers have ignored the importance of tacit practice. However, even though attempts are made to combine the two approaches in legibility studies, there is still a need for the typeface- and graphic design industry to take account of the existing research on legibility related matters. Most of the previous empirical studies into typeface legibility are based on the comparison of different fonts in a retrospective manner after the development has taken place, and not as a part of the design process. These fonts vary so much in overall appearance that it is difficult to say exactly which of the qualities of the individual fonts make them perform as they do in different test situations. Designers have often carried out informal studies, by investigating their own perception of the font at distance, or by asking the opinion of the target group doing the development;

3) Examples of projects combining academic research with the practical knowledge of type design: (a) The ‘Advanced Reading Technology’ group currently working at Microsoft, which includes both typ designers and reading psychologists. (b) The team lead by Herbert Spencer, which carried out work at the Royal College of Art in the sixties and seventies. (c) One of the studies carried out by Poulton (1965) was elsewhere reported by Cheetman & Grimely (1964) one of them a designer; it is a reasonable assumption that Poulton benefited from their knowledge in his work. (d) Cyril Burt (1959) consulted with Stanley Morison and Beatrice Warde in his legibility studies. (e) The study carried out by Barbara Elisabeth Roethlein (1912) was supported by the staff at the American Type Founders Company.

4) Roethlein (1912), Pyke (1926), Tinker & Paterson (1952), Luckiesh & Moss (1942), Zachrisson (1965) are all part of the era of traditional legibility research, with a methodology that compares a range of different typefaces with the aim of discovering the most legible.
however, empirical testing, controlled in a laboratory setting during development, has up until recently, rarely been applied\(^5\). Furthermore, the existing body of knowledge on typeface familiarity consists more or less only of anecdotal references.

---

Scientific and practice orientated approaches are both based on experience; in the former the experience is gained by observation, in the latter it is gained by doing. By combining the two forms of experience in an easy and accessible manner, a central goal of this project is to stimulate interest among the more cautious designers and to inform cognitive psychologists of the existing knowledge of the design community.

5) Publicly available examples of testing during the design process: (a) the work carried out by Jock Kinneir and Margaret Calvert for the design of the British road and motorway signage in 1959. In the first stages of the design process Kinneir and Calvert ran various informal experiments to help determine the stroke thickness and x-height of the characters (Kinneir 1984), after the design was completed, a more controlled study was carried out by the Road Research Laboratory (Christie and Rutley, 1961). (b) In recent years Terminal Design, in collaboration with Meeker & Associates, created the typefaces ClearviewHwy and Rawlinson Roadway for use on American road signs. Various forms of distance tests were carried out during the design development (Garvey et al. 1997; Hawkins et al 1999; Carlson 2003). (c) The typeface family Sassoon by Rosemary Sassoon and Adrian Williams is based on a study of young children’s preferences in reading (Daines 1997). (d) Before finishing the design of Tiresias Screenfont, a simple study was carried out asking participants to rate the typeface in comparison with two other faces (Silver et al 2000). (e) While creating the typeface Sylexiad, Rob Hillier (2006) made an effort to integrate studies of readers’ opinion in the development. (f) Ann M.M. Besseman (2007) is, at the time of writing, in the process of creating a typeface for children with low vision where she incorporates testing in the design process, and finally (g) Karin von Ompteda is working at the Royal College of Art on optimising typefaces for visual impaired adults through experimental studies.
1.3

Defining the meaning of central terms and use of language

The fact that there has been so little crossover between academic legibility research and type design is further demonstrated in the different meanings that each group gives to some rather central terms, and in the different ways in which each group communicates. While an academic argument is ideally motivated by neutrality, the writings presented by designers are often strongly opinionated. Furthermore, in the humanities, the central claim is usually put forward first, and then evidence is put together to support that viewpoint. Argumentation in science starts with the evidence and then a claim is built that matches the evidence.

The designation design has different meanings depending on who is using it. Both practitioners and academics seem to agree that the word describes the process of solving a given problem; however, among designers, this definition is often connected to the creation of some sort of artefact. Yet in empirical research,

6) Two of the earlier influential practitioners who tended to create arguments based on personal views were Daniel Berkeley Updike (1860-1941) and Stanley Morison (1889-1967).
the term more often refers to the planning of an experiment, which has no relation to any artefact that might be part of the test material.

Another and far more problematic overlap is related to the term experimental. In both communities, experimental work is based on try-outs and the search for new discoveries. However, the actual methodology applied appears quite different. When artists and designers produce work of this nature, it is most often motivated by a combination of intuition and theory; conclusions are made based on one’s own rationalizations without any statistical data for backup. In contrast to this, experimental work carried out by academics is placed in a controlled laboratory setting. If the focus is on legibility, the experiment will often involve the participation of a number of readers and a study of their behavior. All stages of the experiment will be planned thoroughly in advance, and with the whole work process subsequently accounted for in such a fashion that others can later repeat the work to prove it right or wrong.

More confusion is found in the terms legibility and readability. One of the early traditional legibility researchers, Pyke (1926), avoided using the term readability altogether. Although aware of the different variables, he chose instead to apply different meanings to the term legibility by differentiating his subject matter into the legibility of (1) letters, (2) words and (3) continuous text. The empirical researcher Miles A. Tinker also preferred one word to cover both topics (Tinker, 1944, 1963). Even today the two terms are often used interchangeably by academics. The reluctance among some academics to apply the term readability is most likely related to the word being frequently used in readability formulas having no association with typography; their purpose is to study factors such as the length and complexity of a sentence and word frequency, and through that measure the complexity of writing.

Nevertheless, at the time of typography scholar Beatrice Warde, and perhaps based on advice of the printing industry, some empirical researchers came to draw distinctions between the two:

When the psychologists reluctantly began using the word ‘readability’ as a term distinct from ‘legibility’, we knew they were realizing that they had erected too narrow a titular fence around their ‘field’ (Warde, 1956, p.55).
As suggested by Warde, this distinction is common among designers. Walter Tracy, defined legibility as follows:

... we want the word to mean the quality of being decipherable and recognisable [...] legibility is the term to use when discussing the clarity of single characters (Tracy, 1986, p.31).

Tracy goes on defining readability:

[It is as] if the columns of a newspaper or magazine or the pages of a book can be read for many minutes at a time without strain or difficulty, then we can say the type has good readability (Tracy, 1986, p.31).

--

To avoid confusion, the term design will in this project be applied solely in relation to the craft; the academic definition will instead be referred to as the procedure. This distinction is not as easily made with the term experimental: the value of the word will depend on the context in which it appears. The designation readability will refer to the level of strain a reader experiences when the eye moves along the line of text; the designation legibility will describe the clarity of letters while influenced by typeface familiarity; the designation visibility will describe the clarity of letters isolated from the influence of typeface familiarity; and the designation familiarity will refer to the collective influence of previous exposure and the level of common letter features.
1.4

The validation of the most essential test methods

Over the years, experimental testing on legibility related matters has been based on a range of different methods and carried out in a range of different ways. None of these methods, however, have avoided criticism for being inadequate for the purpose. An argument often made is that all existing methods for measuring legibility are useless, since reading is such a complex process that no single method will ever show any useful results. This is an argument of essentialism, based on the notion that we must understand legibility fully before studying it. Whittenmore stressed the point by questioning the test methods and the whole issue of legibility in asking:

Do you mean (1) easy to read fast, (2) easy to read at a distance, (3) easy to read in dim light, (4) easy to read when you haven’t your glasses, (5) easy on the brain, (6) not tiring to the eyes, (7) possible to grasp in big gulps of meaning, (8) pleasant to read, (9) inviting to the eye, or (10) something else? (Whittenmore, 1948, p.36)

Another, common criticism is based on the belief that a reader placed in a laboratory setting will always be aware of the action
of reading and, as a result, the possibility of a realistic measurement thereby becomes problematic. Among some practitioners, experimental methodologies are further criticised for being too narrow-minded. Ruari McLean argues, for instance, that:

The findings of most ‘laboratory’ tests of legibility prove, if they prove anything, what suited those people, of that age and sex, at that time of day (Tired? Well-fed? Hungry? In good or bad temper?), in that month, in those conditions (McLean, 1980, p.47).

A related critique is that the human mind is too complicated for any usable information to be extracted during a laboratory test situation. Whenever one problem is tested, a range of other factors will always influence the subject and result in inadequate findings. Arguing against this view, Stanovich (2007) emphasises that studies carried out in a natural environment open up too many uncontrollable variables; instead, he reasons in favour of controlled situations in a laboratory, since this is the only method to separate the many correlating variables that influence everyday life.

The matter of the participant’s behaviour has been of interest to Rosenberg (1965), who carried out a range of investigations concerning the influence on the general test-result in relation to the way the tests are presented to the participants. According to Rosenberg, even in cases where a participant believes that his behaviour is not being directly studied, he will still be concerned to deliver what he believes to be a good performance. This circumstance could potentially influence the speed at which the participant reads a text. However, the counter argument goes that this psychological phenomenon does not influence the test results as long as the study involves comparison. The participant’s wish to deliver a good performance would not differ while reading different passages of text within the same study.

A far more essential problem highlighted by several researchers is the lack of similarities in the results arrived at between different test methods (Tinker & Webster, 1935; Burt, 1959; Lund, 1999). To investigate the matter, Tinker (1944) compared 10 fonts in terms of: visibility from almost clear to dark, percepti-
bility at a distance, speed of reading, and the reader’s opinion of the most legible. He found that there was little agreement between the results of the four test methods. In the measurement from clear to dark, bolder fonts performed better than lighter fonts, hence this study had a lot in common with perceptibility at a distance. He furthermore found that the reader’s opinion was less comparable with speed of reading than the two others, and that readers in general judged fonts that perform well in distance studies to be best for comfortable reading.

Tinker did occasionally publish the specimen sheets he used in his studies (fig.113); however, a large amount of published research results provide no examples of the fonts chosen or of the textual material used. Due to a lack of interest in typography among psychologists, the researchers do not seem to consider the importance of publishing the test material, and the peer reviews do not seem to note the lack of this information as a problem.

Traditional legibility researchers who were interested in typographical matters were looking to discover universal typographic guidelines, an approach that has often been criticised in the design community as an impossible goal – a critique also voiced by Waller (2007), advocating the more target-orientated legibility research of testing fonts for one specific purpose alone.

The various test methods applied in most legibility studies have all emerged from the need to solve problems related to existing methods. The most essential of these can be divided into four categories: continuous reading, search task, threshold, and reader’s opinion.

**Continuous reading**

A problem in many of the studies focused on continuous reading is the countless numbers of variables to be controlled. The isolation of one variable for testing is not an easy matter when dealing with typographically based test material, because variables frequently interact with each other. One of these troubled parameters is related to the leading of the type. The optimal leading of a font with a small x-height is not likely to be the same as the optimal leading of a font with a large x-height. Does this mean that the two fonts should be set with different leading? If
so, how do we ensure that what is being tested is the readability of the font and not also the layout of the type? If font A in reality performs better with a narrow text column than with a wide text column, and the opposite goes for font B, does that mean that font A, tested with a narrow column, is more legible than font B, tested with the same column width, or would font B have delivered a better performance if presented with a wide text column? After discussing these interacting variables, the reading psychologist Linda Reynolds (2007) emphasises the difficulty in generalising from results based on the variation of one factor alone with the others staying constant, and she argues that one single study will never explain all aspects of a problem, but that different methods can contribute in different ways.

**Reading aloud**

One method available for the study of continuous reading is to have the participant read a text aloud and then record the number of errors or the time course afterwards. A dilemma in oral reading is that the situation is unnatural for most adults; as noted by Rayner & Pollatsek (1989) it is difficult to say whether misreadings are based on errors of identification or errors of interpretation or memory. The point being that, when reading aloud, readers often use similar words or restructure the text, which is not to say that they actually encode the text that way. Another dissimilarity is that oral reading causes a higher frequency of fixations and that oral speed of reading is about half that of silent reading (Rayner & Pollatsek, 1989). Morton (1964) further emphasised that due to the 'eye-voice span' when pronouncing a particular word in a reading aloud study, the eye is already further along the line identifying the next words, and as a result only very large performance differences will show up in the test results.

Despite calling attention to some of these issues, Pyke (1926), and later Shaw (1969) applied this method in their work, arguing that even if the errors were dissimilar in oral and silent reading, these differences would probably remain constant from font to font.

7) The term 'eye-voice span' is defined by Levin as, 'the distance the eye is ahead of the voice in reading aloud' (Levin 1979: p.1).
Errors

Errors can also be measured in silent reading. With this method, researchers measure the participant’s level of comprehension after reading a text, an approach closer in style to a normal reading situation. This, however, creates a new range of problematic issues to deal with. How, for instance, do we ensure that the participants all have the same amount of interest in the topics of the text? If they find the topics uninteresting, will that influence their concentration and will their comprehension suffer as a result? Are we in reality testing the participant’s intelligence, or simply their experience with being in a test situation? Furthermore, it has been demonstrated that high frequency words such as ‘the’ are read faster than other three-letter words (O’Regan 1979), and that sentences in the active voice are recognized faster than sentences in the passive voice (Foster & Olbrei, 1973). If the method of an experiment is to compare two different texts set in two different fonts, the individual level of frequency words, and the structure of the text, will likely influence the outcome of the study. However, if all participants were tried on all font conditions, and if test material was counterbalanced between conditions, these issues would not cause problems.

Speed

The method applied by Miles A. Tinker when testing speed of reading, was to have participants read a series of short paragraphs. Each of these paragraphs would contain one phrase towards the end that confused the meaning of the context; the task was then to identify this phrase. More recently, Gugerty et al. (2004) have applied a variation of this study with a forced choice methodology between a right and wrong second sentence following the first.

Another method of studying speed of reading involves a focus on the time it takes to read a normal body text. This can be measured by having the participant read as much as possible of a given text within a time limit, or it can be measured by the time it takes to read a whole text. Both approaches are normally finished off with a study of comprehension accuracy. Legibility researchers often stress the importance of comprehension in these studies, as Tinker emphasised: “‘Reading’ without understanding is not reading” (Tinker, 1964, p.22).
Several early legibility researchers (Pyke, 1926; Luckiesh & Moss, 1940) argued that the assumption that speed of reading increases as legibility increases might be a mistake. Their point was that in some situations, when reading a highly legible text, the participant performs the task with less effort instead of increasing the rate of reading; a form of behaviour that does not show in a speed of reading test situation. Chaparro and colleagues (Chaparro et al., 2004), further found that in a comparison between text set with large margins and text set with no margins, that reading speed in the large margin text decreased and comprehension increased.

Another central criticism of the speed of reading method is the frequent lack of significance in the measured time differences between the fonts tested. Tinker (1968) was aware of this issue, noting that when one font is read faster than another it means that it is easier to read and therefore legibility must be a key factor. He nevertheless stressed the point that, if no significant difference is found between two fonts, it does not necessarily mean that there is no difference in legibility; the simple problem is that the method is not sensitive enough to detect the difference.

The search task
This method involves a visual search with no measurement of comprehension. Participants are asked to locate spelling errors or specific words; Poulton recommended the method as ‘quick and easy to use’ (Poulton, 1968, p.73). The search task method was applied in studies carried out by the research unit working at the Royal College of Art in the sixties and seventies. Their choice of method was motivated by a disregard for the comprehension testing. The researchers viewed the comprehension method as having too many unaccountable variables unrelated to the visual qualities of the text (Reynolds, 2007).

It is however possible that since participants know beforehand what they are looking for, the interrelation between the top-down process of reading the word, and the bottom-up processes of reading the single letter, would be more top-down driven than what is normal when reading without searching for a specific target (see more in Chapter 1.4). This concern is confirmed by a distance study of traffic signage carried out by
Garvey and colleagues (1997), finding no difference in performance between all-uppercase words and initial capitalized lowercase words, when words were unknown to participants beforehand. However, when testing the recognition of already known words, the initial capitalized lowercase words delivered a better performance than the all-uppercase words. The researchers speculated that this was caused by the reader’s mental image of place-names often being initial capitalized lowercase words, which makes the cognitive task of matching easier with initial capitalized lowercase words than it does with uppercase.

**Visual accuracy threshold**

The focus of this method is on letter and word identification with comprehension as a non-priority. When discussing the validation of legibility based studies, supporters argue that skill based behaviour – in this context, identification of shapes – does not tend to be influenced by a test situation, as demonstrated by the fact that optometrists are able to make rather accurate measurements of the eye in a laboratory setting (Sheedy et al., 2005). The argument is that word recognition happens on an automatic level, and therefore is not influenced by surroundings. The theory is further supported by the Stroop effect (Stroop, 1935) showing that people have trouble turning off the automatic processing of word recognition (fig.1). People recognize words even in cases where they are supposed to ignore them. The theory is that this happens because the mind automatically perceives the meaning of the word, an action independent of the task of identifying the ink colour, and therefore an action that would not vary according to setting.

**Variable distance**

In a distance threshold study, researchers investigate the relationship between type size and distance to the eye, defined as the visual angle. This is a relevant method for research into perception of signage in general. The various tests carried out in relation to the development of the new typefaces for American road signs are a good example of this (Garvey et al. 1997; Hawkins)

---

8) The all-uppercase words were set in the font Standard Highway Series D, and the lowercase words were set in an early version of a Clearview font.
et al. 1999). It is furthermore a method easily and informally applied by designers, as did Jean François Porchez when he convinced the editor of the French newspaper *Le Monde* to replace Times New Roman with his new designs (Suzuki, 2004).

Focusing mostly on print material in his work, Tinker (1964) noted that this kind of study shows no agreement with speed of reading tests. Like Herbert Spencer (1968), Tinker endorsed the usefulness of the method in so far as it can estimate the relative legibility of letters. Zachrisson, on the other hand, only saw the value of the method in situations where 'the concept of distance is of importance to the problem under study' (Zachrisson, 1965, p.67); a view somewhat supported by the influential printer Daniel Berkeley Updike, who, in 1941, expressed his dislike of a study which applied this method, since: 'a book is held at only a number of inches from the eye' (Updike, 2002, p.46). He went on to note that the experimenters had most likely chosen the method based on experiences with the oculist, and, as a consequence, mistook visibility for legibility. This criticism leads to the work of optometrist James E. Sheedy and colleagues (2005), in which the distance threshold method was employed. Here the argument goes that if a font is identifiable at a small point size, the larger the difference in sizes between the lowest identifiable size and the size applied for reading, the higher the legibility. This description is quite close to what Updike would define as visibility. However, Sheedy points out that although it seems logical that this matter would improve reading performance, the fact has yet to be established. While many issues can be raised with regards the validation of this method, when applied to the study of distance accuracy of a signage typeface, the method has rarely been questioned.

**Short exposure**

In a short exposure study, participants will be exposed to the stimulus for a brief period of time. After a rapid exposure, at a length short enough to prohibit the eyes of the participant to move from one fixation to another, the participant is asked to identify the material shown. The method is useful for studying the legibility of the individual characters or words. However, because of the single fixation, it has been claimed that the method varies too much from continuous reading, and therefore the re-
The relationship between the two should be looked at with caution (see Spencer, 1968; Tinker, 1964).

On the other hand, one could argue a similarity in the two situations based on the fact that the eyes are relatively stable in carrying out the fixation, both in the short exposure and in continuous reading conditions. A reasonable assumption is that, because of the similarity in stimulus to the eye, the perceptual processes would also be the same. Yet according to McConkie (1983) this is not always the case. It is frequently seen that humans process the same information in different ways depending on the task they set out to perform. In the short exposure method, the participant prepares him or herself for the task of perceiving the material in one fixation. In continuous reading, the fixation is a part of the whole, and is not seen as one specific task. McConkie does not argue against the use of the short exposure method; instead he indicates that one test method cannot cover every issue, and therefore that several methods must be employed to cover all aspects of the reading process.

Readers’ preferences

All previous methods are performance based and objective. Studies into readers’ preferences, however, focus on the subjective opinion of the participant. Asking the preferences of the target group can generate helpful information when studying whether a typeface appeals to the public or not. Yet it is not the most useful method when studying visibility, since the preference ranking of a certain font will likely be influenced by the reader’s personal view on aesthetics. Another aspect, discussed above, is the fact that most participants wish to please the experimenter. While it is difficult to improve one’s own performance in reading in speed and accuracy tests, it is quite easy to answer preference questions favourably if you can guess what condition the experimenter cares about. The main purpose of this method lies in the elaboration of reader’s opinion and motivations, with a rather low value when performance is of interest.
Three categories of the four presented rely on task-based behaviour. One defines legibility as comprehension and speed, another as the speed of which a participant is able to find a certain item within a range of others, and the last category leaves it to the reader to rank the test material. Testing these parameters demands a certain level of mental awareness of the tasks carried out. In contrast to this, the category of visual accuracy thresholds primarily relies on the perception system. As a consequence, the participants tend not to be as easily influenced by the unusual situation of the laboratory setting. However, if the aim is to investigate a font’s ability to function in continuous reading, this methodology is not ideal. Unfortunately, very few of the methods described are without faults. Results reached by any one of the methods mentioned would therefore benefit from being considered in combination with others and analysed with a critical mind.
1.5 Perception of letters

Defining what actually happens in the perception of letters and words is still a rather controversial subject. Cognitive psychologists have come up with a number of theories trying to understand the act of reading. These theories range from the notion that we perceive the words as wholes without recognizing the individual letters (Huey, 1908), to the idea that reading is based on a letter by letter recognition process (Sperling, 1963; Gough, 1972).

Although we have yet to understand completely how the brain works during reading, the aim of this chapter is to present the arguments that lead to the model of the most plausible hypothesis regarding the perception of letters and words.

Letter identification

The two main theories on the subject of letter identification are the template-matching and the feature-comparison theories; the former defined as a holistic approach where characters are perceived as a whole, and the latter being a more analytic approach dividing the characters into different elements in the perception process.
The basic idea of the template-matching theory is that for each letter of the alphabet, the brain has stored a basic template of the letterforms. As we perceive a new shape, the brain goes through a series of templates to find the one that matches the best. This is a logical assumption when trying to understand the process of letter perception, and also the idea of the renowned type designer Adrian Frutiger, who compares the function of reading to a keyhole and its key, where the reader locates the basic skeleton form of the letter that then fits like a key into the keyhole in the identification (Hunziker, 1998).

However, the main problem with this theory is how it is able to explain the wide variations in typefaces and handwriting that we are capable of perceiving. Does it mean that the brain has a separate template for a flamboyant ‘A’ and a simple Sans Serif ‘A’, or for all variations of handwriting? Even if the brain has some form of clean-up process of the shapes, it seems doubtful that a system like this can decide which part of a character shape is essential and which is not (Naus & Shillman, 1976; Crowder 1982; Rayner & Pollatsek, 1989; Underwood & Batt, 1996; Smith, 2004); this problem leads us to the feature-comparison theory.

Instead of perceiving the whole character, the basic idea behind the feature-comparison theory is that the brain decodes the different features of the character individually. The analytic process is based on a perception of the characters as a range of small features where the elements are put together until a stage of identification occurs (fig.2).

Figure 2: The internal relationship of letters in the feature-comparison theory illustrated in Neue Helvetica
An argument in favour of the feature-comparison theory is based on the work of Hubel and Wiesel (1962) involving the study of the visual system of a cat. By projecting different forms of patterns into different regions of the retina, Hubel & Wiesel demonstrated that the cortical cell in the visual system of the cat fired differently according to the type of stimulus being processed as horizontal, vertical, or curved shapes. There are obviously various differences between the vision of a cat and a human, however it is commonly accepted that this identification process of lines and curves in the visual cortex is rather similar between the two species (Crowder, 1982; Rayner & Pollatsek, 1989).

Another finding supporting the feature-comparison theory is based on a study by Neisser (1967), who discovered that participants found it easier with a search task to locate the letter ‘z’ in a group of visually unrelated characters (ODUGQRC), than in a group of related characters (IVMXEW) (fig.3).

![Figure 3. Lists for visual searching. The target is ‘z’ in both lists (Neisser, 1967, p.70).](image-url)

<table>
<thead>
<tr>
<th>ODUGQR</th>
<th>IVMXEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>QCUDGO</td>
<td>EWVMIX</td>
</tr>
<tr>
<td>CQOQRD</td>
<td>EXWMI</td>
</tr>
<tr>
<td>QUGCDR</td>
<td>IXEMWV</td>
</tr>
<tr>
<td>URDGQO</td>
<td>VXWMEI</td>
</tr>
<tr>
<td>GRUQDO</td>
<td>MXVEWI</td>
</tr>
<tr>
<td>DUZGRO</td>
<td>XVWMEI</td>
</tr>
<tr>
<td>UCGROD</td>
<td>MWXVIE</td>
</tr>
<tr>
<td>DJQCGU</td>
<td>VIMEXW</td>
</tr>
<tr>
<td>QDOCGU</td>
<td>EXVWIM</td>
</tr>
<tr>
<td>CGUROQ</td>
<td>VWMIEX</td>
</tr>
<tr>
<td>OCDURQ</td>
<td>VMWIEX</td>
</tr>
<tr>
<td>UOCQGD</td>
<td>XVWMEI</td>
</tr>
<tr>
<td>RGQCOU</td>
<td>XMEWIV</td>
</tr>
<tr>
<td>GRUDQO</td>
<td>MXIVEW</td>
</tr>
<tr>
<td>OCRUQ</td>
<td>VEWMI</td>
</tr>
<tr>
<td>DUOCQG</td>
<td>IVWMEX</td>
</tr>
<tr>
<td>CGRDQU</td>
<td>IEVMWX</td>
</tr>
<tr>
<td>UDRCOQ</td>
<td>WVZMXE</td>
</tr>
<tr>
<td>GQCORU</td>
<td>XEMIWX</td>
</tr>
<tr>
<td>GOQUCD</td>
<td>WXIMEV</td>
</tr>
<tr>
<td>URDCGO</td>
<td>EMWIVX</td>
</tr>
<tr>
<td>GODQGC</td>
<td>IVEMXW</td>
</tr>
</tbody>
</table>
The template-matching theory would suggest that the results should show no difference between the unrelated and related character groups. Since the eye is searching for one template, the shapes of the surrounding templates would be of no importance. However if we analyze the results by applying the feature-comparison theory, searching for the letter ‘z’ in a related character group means searching among a range of similar features, which would be a plausible reason for the slowing down of the search process.

However, support for the feature-comparison theory is not fully substantiated by a study carried out by Pritchard (1961). Pritchard discovered that while fixating an image to the retina, in such a way that when the eye moved the image moved simultaneously with the eye, the image would eventually disappear. Pritchard found that complex stimuli sometimes disappeared and reappeared as a whole, and sometimes vanished in fragments. This suggests some level of presence for both the template-matching and the feature-comparison theories, and that the two theories of perception are interrelated.

Single letter and word superiority effects

In 1886, James McKeen Cattell showed that, after a short exposure, participants were more likely to identify single words than single letters. This discovery is defined as the word superiority effect. Reicher (1969) later recreated the experiment with a few adjustments, pointing out that in the earlier experiment it was possible that participants did not perceive all letters of a word, but were still able to guess the word based on parts. Reicher therefore changed his experiments to a forced-choice between two alternative single letters shown after the presentation of the stimulus (fig.4). He did this in such a way that the two alternative choices would both make up a word in connection with parts of the stimulus word. Reicher found that letters in words were more accurately recognised than both single letters and nonsense words (see Wheeler 1970 for similar results). A hypothesis based on these findings is that reading is based on a long-term memory of words and word patterns.

However, other studies show that this is not exactly the case. Adding pronounceable nonwords (pseudo words) such as mave, or reet to experiments based on Reicher’s method, a number of researchers (Aderman & Smith, 1971; Baron & Thurston, 1973;
found that these words are – in most cases – far better recognized than unpronounceable nonwords (nonsense words) such as ‘ftgy’, or ‘ojhl’, indicating that the word superiority effect is a result of letter combinations rather than a result of familiar word patterns. Reicher (1969), Wheeler (1970), McClelland & Johnston (1977), and partly Massaro & Klitzke (1979), all demonstrated that words are more easily recognized than single letters.

Sheedy and colleagues (2005), on the other hand have established that the opposite is the case. In a study applying Verdana as the stimuli, they found that single lowercase letters were 10% to 20% more legible than lowercase words. However, contrary to the other studies that were all based on short exposure methods, Sheedy’s investigation was based on a distance threshold study. Sheedy recognizes that ‘crowding’ may be influencing the findings; an effect sometimes also referred to as ‘contour interaction’, where the viewer finds it difficult to identify a letter embedded in other letters, due to letters next to each other interfering, and in that way lowering recognition. Research into the crowding phenomenon has been a popular subject in recent years. A look through some of these studies suggests that the phenomenon is most present in the parafoveal and peripheral areas (Chung et al., 1998; Pelli et al., 2007), and – as in the Sheedy study – in the fovea when the material is viewed at a distance (Hess et al., 2000; Liu & Arditi, 2001). This fact has long been known in the art world, where pointillist painters have explored the effect by showing that different colour dots merge when perceived from afar. According to typodesigner Jock Kinneir (1978, 1980) typefaces for distance viewing should for that reason have wider letter spacing than that which is usual in continuous text.

Only a few of the studies presented here have an even minor concern with the individual legibility of characters – a major area in the traditional legibility research, however not of high interest in cognitive psychology. As will be discussed thoroughly

9) Massaro & Klitzke (1979) found that words were more easily recognized than single letters at short time intervals between the presentation of the stimulus and the forced-choice test, while the opposite was the case with a long interval.

10) The study by Sheedy and colleagues considered the influence of the individual letter by excluding the lowercase characters g, l, m, t, w, i and j due to ‘unique legibility traits’ (Sheedy et al., 2005, p.800).
in a later chapter, traditional legibility research has shown that a number of characters have a higher tendency to be misread than others, and some characters even tend to be completely overlooked. Acknowledging this issue ought to be a central factor in any selection of words and letters in stimulus material. However, it is also apparent that none of the researchers have any typography knowledge. A striking fact is that even when being thorough on all other aspects of the methods applied, only a few of the studies contain any information on the typeface applied or discuss its influence on the final result, and none of the papers discuss matters such as spacing, stroke contrast and weight. A number of the studies do not even specify whether the stimulus material is created out of upper or lowercase characters. This leaves us with a string of variables unanswered. If the typeface applied is Courier (as it probably was in many of the studies from the 70s and 80s), it is likely to influence the difference in perception of single letters and words. It is possible that the large serifs would alter the shape of characters; however, when seen as part of a word in the context of other characters with similar serifs, the serifs would not seem as dominant.

Although the Sheedy distance study showed otherwise, the collective data of the short exposure investigations demonstrate that letters in words and pseudowords are, in most cases, perceived more accurately than single letters. However as we shall see, this does not necessarily mean that we read by word wholes alone.

Word wholes
Many internet users have encountered a circulating text referring to a research project which found that ‘it deosn’t mtaer waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteres are at the rghit pclae’, the text further concludes that "This is bcuseae we do not raed ervey lteter by it slef but the wrod as a wlohe". Testing this jumbled word effect, Rayner and colleagues (2006) found that reading speed in general slowed down when letters were transposed.

11) In the complete text, the research is asserted to have been carried out at Cambridge University, however according to Matt Davis (2003) working at 'Cognition and Brain Sciences Unite' at Cambridge, this is not the case. Davis identifies the research as originating in a PhD thesis written at Nottingham University by Graham Rawlinson in 1976 (for summary see Rawlinson 2003).
Matt Davis (2003) lists a range of elements that influences the readability of the text in question, arguing that the swapping of letters is much easier to read when (1) two neighbouring letters are switched (as in porbelm for problem) than it is when the letters originated further apart in the word (as in pborlem); that (2) none of the swapped letters create new words or new sounds (toatl instead of ttaol for the word total); (3) that the content of the text is reasonably predictable, making it possible to guess a word from its surroundings, and (4) that all the function words (the, be, and, you etc.) stay the same preserving the grammatical structure. Finally he points out that if reading is based on word wholes alone – as the text claims – then the shifting of ascending and descending letters would disrupt the word shape and by that, the identification of the word. One can further assume that the phenomenon will be less successful when applied to languages with many compound words, such as the Scandinavian, German, and Dutch.

If we do read by word wholes, then words set in mIxEd cAsEs would slow down the reading dramatically. Smith and colleagues (1969) found that both in a reading aloud and in a search task study, words set in mixed cases where the characters were kept in their original x or cap-height sizes did not perform well. However, it was also found that words set in mixed cases, where the height of the upper and lowercase letters were adjusted to the same size, performed equally with words set only in uppercase or lowercase letters. Coltheart & Freeman (1974) later criticized their methods of speed of oral reading and search task for not being able to show any existing difference in performance. In a short exposure study, Coltheart & Freeman themselves found the worst performance was with the adjusted mixed case words.

Adams (1979) emphasised, however, that this inferiority of the mixed case words is more likely related to familiarity than to word wholes, and proved the argument by showing that the performance of mixed case pseudowords and mixed case regular words were equally slowed. Were it true that we read by word wholes, the performance difference between familiar regular words changed to mixed case words would have been radically different from the corresponding pseudowords, since the pseudowords do not contain any familiar word shapes, either when set in normal case or in mixed case.
In another study carried out by Haber & Schindler (1981), participants were asked, in a proofreading study, to locate misspellings. Haber & Schindler found that misspellings that changed the overall shape of a word (test to tesc), were more likely to be detected than misspellings that were consistent with word pattern (test to tesf). Paap and colleagues (1984) later showed that this phenomenon is, in fact, more closely related to a similarity of letter shapes than to word shapes.

Parts, wholes & context
Most of the studies reviewed so far fit the basic ideas of the Parallel Letter Recognition (PLR) model (McClelland & Johnston, 1977; Rayner & Pollatsek, 1989; Larson, 2004) (fig.5). The model contains three basic levels: the first being the feature detection level. As earlier described, the process at this stage is to recognize the features of the individual letters, such as horizontal, vertical, curved and diagonal lines. This information is then sent to the letter detector level. If an ’o’ is part of the stimulus material, the letter detectors for ’o’ would be active in combination with letter detectors for other related shapes such as ’c’ and ’e’. The task for the letter detectors is to locate the letter with the greatest amount of common features identical with the information received from the feature detection level. The final level involves the word detectors functioning in the same way as the letter detectors by identifying the features (letters) and combining them into words. What further happens on the word

Figure 5. The Parallel Letter Recognition (PLR) model, based on the model presented by Rayner & Pollatsek (1989), a simplified version of the ‘Activation-Verification model’ by Paap and colleagues (1982).

A variation of the model is identified as The interactive-activation model by McClelland & Rumelhart (1981); and Rumelhart & McClelland (1982).
detector level is not completely identified (see McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982; Paap & Noel, 1991). However, it appears that a second process takes place on the word detector level, consisting of a top-down input of some kind of lexical stimulus on context, word wholes and word parts. This operation then proceeds further down to the letter level in a parallel process.

This top-down and bottom-up processing of the PLR model explains the word superiority effect. Whereas single letters have to be identified exclusively by information received from the letter detectors, words receive information from both the letter detectors and the word detectors, and therefore presumably words will have a higher recognition rate than single letters (McClelland & Rumelhart, 1981; Rayner & Pollatsek, 1989). When the perceived word is not identified in the word lexicon, we spell our way through the word, using only the letter level. If a couple of letters are unidentifiable, the collaboration between the word lexicon and the letter lexicon will be able to identify the word. The model further explains the jumbled word effect; it seems reasonable that as long as there is no phonetic confusion, the collaboration between the predictability delivered by the top-down process, and the letter identification in the bottom-up process is capable of identifying swapped characters as long as they are not placed too far apart.

A recent study carried out by Pelli & Tillman (2007), takes a different angle to the matter of internal relationship between the different processes influencing reading. The researchers set out to isolate the three mental processes of: letter-by-letter (L), word-wholes (W), and sentence-context (S) recognition.

<table>
<thead>
<tr>
<th>knock-out (S)</th>
<th>Contribute others. the of Reading measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>knock-out (W)</td>
<td>This text AIrEtnEEs iN CaSe.</td>
</tr>
<tr>
<td>knock-out (L)</td>
<td>This sartcrec bes lctfan suhsflufas.</td>
</tr>
</tbody>
</table>

By measuring reading speed using rapid serial visual presentation, and oral and silent reading from printed pages, the manipulations shown in figure 6 were tested both one at a time and in combinations. Within the three knock-outs, the letter process
(L) came across as the strongest accounting for 62% of the adult reading rate, the sentence (S) process came in second accounting for 22%, with the weakest knock-out being the word (W) process accounting for 16%. The study further found that the three processes all work together. The researchers argued that if the three processes were working independently of each other, when all were at work collectively, the work carried out by the two weaker processes W and S, would be overruled by the work of the strong process L. This would mean that if W and S were operating alone, their performance would be better than when L is present. That was not the case, and the researchers concluded that the three processes operate as individuals coordinating the distribution of tasks amongst each other.

Combining the study of Pelli & Tillman with the ideas of the PLR model, gives a good indication of the different kinds of operations that takes place in the reading process. It appears that the functions of letter, word, and context detectors, collaborate with each other by approaching the reading matter from different angles. Although highly dependent on the other detectors, the function of identifying the individual letter comes across as the strongest of the forces.
1.6

The internal relation of visibility and familiarity

Designers have long had the idea that familiarity has a role to play in the reading process. In the legibility debate, however, the visibility and familiarity of a typeface have yet to be clearly defined as separate parameters. As a consequence, the discussion tends to stagnate around the undefined influence of familiarity without any real empirical verification.

The cognitive aspect of separation in visual perception

The way we recognize a specific object is by mentally comparing its features with our memory of prior exposures to similar features. The identification will consequently be based on a previous familiarization with a given shape. It could therefore be argued that visual and memory based processing must originate in the same operation. However, that is not necessarily the case. Some patients with brain damage show that loss of object recognition can actually occur independently from the loss of sight (Madsen 2006).

Hochberg (1968) found that when comparing two words set vertically – giving participants the task of identifying the words
as either the same or different – in situations where the words were set so close that they could be perceived simultaneously, the speed of identification was unaffected by them being either known or meaningless words, or being mirrored or normal type (fig.7). However the one condition that slowed down performance was when one word was set in uppercase and the other word in lowercase letters. In a second study, Hochberg separated the words so that they had to be perceived in two fixations. Now the mirrored and meaningless words took longer to identify than the known words. It further showed that there was no influence on results whether the two words for comparison were set in the same or in different cases.

This extra operation of naming the word appears to influence all trial conditions where words are perceived in separate fixations. Based on these findings, Hochberg concluded that the perceptual process consists of two separable components, one being ‘the features glimpsed in momentary glances’, the other being ‘the integrative schematic map into which those features are fitted’ (Hochberg, 1968, p.330).

This conclusion corresponds with that of another same-different study presented by Posner & Mitchell (1967). The experiment contained two separate instructions, one with the task of identifying the physical shape of letters as being the same or different, the other with the task of identifying the name of letters to be
the same or different. The study showed that when participants were asked to identify the physical shape of pairs like A-A (same) and A-a (different) the reaction time was quite similar between same and different; however, when the task was to identify the name of letters, the pair A-A was identified quicker than the pair A-a (fig.8). These findings suggest that visual identification is enough to perform the task in the A-A comparison, whereas when the task is to name the letters A-a, the reader applies a second process of comparing the stimulus to a mental list of familiar features, before being able to identify the letter and then naming it as being the same. It further appears that in the identification of the physical shape of letters, the second process does not automatically have to take place: it is not essential to know what you are comparing in order to identify whether two objects are the same or different. This theory is further confirmed by the fact that participants in Neisser’s search task (discussed in Chapter 1.5), reported not ‘seeing’ the individual letters that were rejected in the search for the target letter.

In relation to legibility, the two separate operations can be interpreted as follows: the first operation is the instant perception, a parameter not influenced by learning and purely based on a visual registration of forms. The second parameter is related to our prior knowledge and stored lexical information on letterforms, and the various versions of letterforms experienced through the different typefaces we have encountered. Familiarity is therefore a central factor in this second operation, and non-essential to the first operation, which is merely driven by the visibility of the shapes.

<table>
<thead>
<tr>
<th>Instructions</th>
<th>Same</th>
<th>Different</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Physical identity (AA)</td>
<td>Name identity (Aa)</td>
</tr>
<tr>
<td>Same physical shape?</td>
<td>428</td>
<td></td>
</tr>
<tr>
<td>Same name?</td>
<td>452 523</td>
<td>461 556</td>
</tr>
</tbody>
</table>
‘Readers read best what they read most’

The Emigre type designer Zuzana Licko famously stated that, ‘Readers read best what they read most’ (Licko, 1990, p.12). In this first part of a longer argument, Licko followed other designers such as William Addison Dwiggins, Eric Gill, Adrian Frutiger, Herbert Bayer and Walter Tracy, their point being that typefaces are only legible if they are familiar to the reader. According to Licko, when the typeface Times Roman first came out, readers were not used to reading it, and it is only because of its frequent use that it has become legible today. Licko went on to speculate whether, in two hundred years, her own typefaces would be viewed as being legible (Licko, 1990). Voicing this opinion, Licko advocated a rather passive approach, in which the designer has little control over the outcome, since the issue of legibility is something that would only become apparent after the typeface had been out on the market for a number of years. Several traditional legibility researchers have also noted the influence of familiarity on the reading experience (Burt, 1960; Tinker & Paterson, 1932; Pyke, 1926), indicating that typography would benefit from not changing style too much. This view seems to suggest that readers would be best served by reading a very small number of typefaces and that new typefaces will not serve the reader.

In the second part of her argument, Licko defined legibility as being a ‘dynamic process, as readers’ habits are ever changing’ (Licko, 1990, p.12). This particular view chimes with the more active approaches advocated by Eric Gill (1936), William Addison Dwiggins (1947) and – less surprisingly – with the view of the Bauhaus typographer Herbert Bayer (1967). These earlier craftsmen all believed that the designer is in a position to control the legibility of the type, and they stressed the importance of

---

12) ‘The kind of letter to which a reader is accustomed is the best kind of letter for the reader’ (Dwiggins, [1919] 1947, p.40).

13) ‘Legibility, in practice, amounts simply to what one is accustomed to’ (Gill, 1936, p.44).

14) ‘I think legibility is solely a matter of habit, and speed in reading depends not so much on the speed of the eye than on that of the mind’ (Adrian Frutiger in Eurographic Press Interview, 1962, p.260).

15) ‘What one is used to is always easier to read than what one is not used to’ (Bayer, [1958-60] 1967, p.80).

16) ‘...the common element in it all is the familiar alphabet’ (Tracy, 1988, p.84).

17) This example would not have pleased Stanley Morison, the initiator of Times New Roman, who strongly believed that 'for a new fount to be successful, it has to be so good that only very few recognize its novelty' (Morison, 1930, p.63).
pushing the less functional familiar features towards more functional ones. As Dwiggins put it: ‘If the reader is used to bad design, he must be led to accustom himself to better design’ (Dwiggins, 1947, p.40). Frederic Goudy (1940) partly supported the notion, but he also argued that any changes made should be small and within strict limits, and he emphasised the essential in sticking with already defined and accepted letterforms.

---

Gill, Dwiggins and Goudy all left behind a legacy of high craftsmanship and a deep interest in the reader’s well-being. Having seemingly no knowledge of the theories put forward in cognitive psychology, these designers recognized familiarity at the same time as they argued for changes to less legible characters towards features of a higher visibility. This suggests that, in their work, they focused on visibility and familiarity simultaneously, and that they also recognized the two processes as being separate aspects of their own designs. They furthermore saw the matter of familiarity as being something they could take an active part in by slowly educating the reader towards typefaces of a more functional nature.
Section 2

VISIBILITY
2.1

The visibility of typefaces for different needs

Some typefaces are designed for a broader usage, others for a particular medium. The specific requirements of a low-resolution screen typeface might be different from the requirements of a print typeface, which again might differ from those of display typefaces. It is likely that, depending on application, certain character features can influence performance – a fact that has not been of much concern in the academic world, although it has often been expressed by type designers. Gerard Unger, the designer of a range of high profile typefaces, including one for Dutch road signs, and several for small print and newspapers, emphasises that the reading of few words at a distance is fundamentally different from the reading of continuous text (Walters & Oliver 2001). This view is further supported by other designers of signage typefaces such as Adrian Frutiger (Nicolay 2004) and David Kindersley (1960). Sumner Stone, however, points out that because of the distance involved in viewing signage, a ‘letter which is four inches tall appears to the reader to be only a 6 pt character when viewed from the appropriate distance’ (Berry, 2006, p.27). This results in a similar visual angle, causing a number of the issues concerning the two categories to be identical, yet as previously discussed (Chapter 1.5) the crowding
phenomenon appears to be more dominant in the fovea for typefaces at a distance than for typefaces up close.

To make the reader unaware of the action of reading continuous text, the transaction between the content and the reader should be as easy as possible. This case is not necessarily the same for signage typefaces in which the process could be more closely related to the decoding of shapes, and therefore the individuality between the characters enhances perception. Furthermore, when a running text is set in small sizes, more characters are perceived simultaneously both in foveal and parafoveal areas. Discussing the importance of the minor adjustments to a typeface, Chauncey H. Griffith compared the individual pieces of type to the threads woven into a textile. Viewed alone, the exact colour would be difficult to identify, however when woven into a cloth the colour would appear obvious (S. Carter, 2002). If a font has a high internal irregularity, when shown small, the diversity will appear domineering since a great number of irregular elements will be present to the reader all at once. On the other hand, if the same font is shown in larger sizes – meaning fewer letters in the foveal vision – it will present the reader with fewer irregular elements in each fixation, and consequently a higher differentiation level of the font is acceptable to the reader. As a result, extreme internal variation is not necessarily the most suitable for fonts designed for running text in small sizes. This matter is illustrated in the typeface Info by Spiekermann and Schäfer, giving differentiation extra focus in the signage version, compared to the text versions of the typeface (fig.9).

Continuous reading

In continuous reading, the eye jumps across the line in rapid saccade motions. Between the saccades, the eyes stop and pause in fixations. The retina has two types of receptors – rods and cones – and these serve quite different purposes. The foveal area consists almost entirely of cones, while moving away from the fovea, the numbers of cones decreases at the same time as the numbers of rods increases. Whereas the rods have the ability of detecting movement and are more sensitive in low light situations, the cones allow us to process details and see sharpness. Thus, the further away from the fovea, the more difficult it is to identify an object. Studies show that not only the foveal area, but also the parafoveal area plays an important role in

Figure 9. The typefaces InfoText and InfoDisplay by Ole Schäfer and Erik Spiekermann.
continuous reading. O’Regan (1979) found that the length of the next word influences the lengths of saccades: long words created longer saccades; short words shorter saccades. He also demonstrated that the word ‘the’ received remarkable fewer fixations by English speaking readers, than other three-letter words. Other studies (Rayner, 1978; Rayner, McConkie & Ehrlich, 1978) support the notion in finding that by changing the word as soon as a saccade moved towards the word, the greater the similarity between the visual patterns of the original word and the replacement word, the sooner the word was identified by participants. Based on these findings, it seems likely that the parafoveal area is of great importance to the continuous reading process, and therefore of importance to text typefaces.

Due to the eye movement from left to right in continuous reading of the Latin alphabet, type designer Jean-François Porchez (1998) emphasises the horizontal movement of text typefaces by focusing on the ends of the character strokes and the angles of the axis, his point being that since most characters have a structure that naturally emphasizes the vertical line, a horizontal emphasis would help the eye movement and enhance the readability of the face (Suzuki 2004). Another way of accentuating the horizontal flow can be seen in the typeface Balance (fig.10), designed by Evert Bloemsma, who was inspired by the typeface Antique Olive, which applies an inverted stress to the characters with the horizontal parts being heavier than the vertical parts. According to Bloemsma, the inverted stress helps to lead the eye along and compensates for the lack of serifs (Middendorp, 2004, 2006d).

A different way of emphasising the direction of reading is seen in the regular styles of Quadraat (fig.11) by Fred Smeijers (2006) and Trinié (fig.12) by Bram de Does (Middendorp, 2004). Inspired by early manuscript letters, both typefaces have a subtle slanting of characters in the reading direction, and the slightly longer serifs to the right of Trinié aim at a stronger overall image.

Designers, often accentuate the significance of the serif in relation to the continuous reading of printed material. The reality is that we do not at this point know the real function of the serifs, a lack of knowledge that leaves us only with assumptions. One argument has it that the horizontal strokes of the serifs help the eye to stay on the line of text and not wander off. Ac-
cording to Middendorp (2004) the serifs of the typeface Avance (fig.13) by Evert Bloemsma, do just that by guiding the eye in a forward direction with solid shapes on the top left of the stems and on the bottom right. Unger (2007a) furthermore speculates that serifs on the extremes of ascending and descending characters enhance the shape of the words in the parafoveal vision. Reviewing 28 studies on the subject of relative legibility of Sans Serif and Serif typefaces, Lund (1999) finds no valid proof in favour of either one of the two.

Among other things, William Addison Dwiggins enjoyed creating marionettes. In 1937, he made a discovery while cutting a head for a new doll. Dwiggins found that to successfully carry the facial expressions of a young girl to observers at the back of the room, the otherwise soft features should be cut as sharp edges (fig.14). At a distance these exaggerated features would appear just as gentle as they were originally intended. Dwiggins later transferred this knowledge into his text typefaces by sharpening the character edges (fig.15). For him, the discovery was a way to trick the eye into seeing nonexistent curves in objects of reduced sizes while enhancing the character’s features. Named after the marionettes, he called this observation his m-formula (see Unger, 1981; Wardle, 2000). Discussing Dwiggins’ theory, Gerard Unger supports the idea that the influence distance and light have on marionettes can be transferred to text of small sizes, yet Unger emphasises that this technique will be most effective in news letters of 7 point or less (Unger, 1981). Based on the notion that text in small sizes and text at distance have a similar visual angle, and since the m-formula originates in a situation involving distance, one can assume that the approach would be useful when applied to type designed for distance as well.

**Type at a distance**

The central issue of keeping the eye along the line of text in continuous reading, is of less importance in relation to signage typefaces, since signs mostly consist of single words and rarely of long paragraphs. While both continuous text and road signage are mostly read in a frontal position, other signs will often need to be read at a more acute angle. Per Mollerup (2005) stresses that if letters like ‘a’, ‘b’, ‘d’, and ‘e’ become too narrow in a condensed signage typeface, when viewed at an angle...
the counters would be obstructed, with a consequently lower legibility - a view supported by Jock Kinneir, who emphasised that condensed letters are less legible than wide letters, and that the option of extra height in the condensed letters would not help matters as long as the counters of the narrow letters merge (Kinneir, 1980).

In an investigation into legibility reported by Robert Waller (2007), a comparison was made between typefaces for signage. The study found Frutiger Bold to be more legible than Frutiger Roman, BAA Sign, Stemple Garamond Italic and Vialog (fig.16). The typeface Vialog appeared to be less legible than both of the Frutiger weights and BAA Sign. The fact that Vialog, which is designed for high legibility, performed so poorly is rather interesting. Waller speculates that the width of the fonts might have influenced the results, since the broadest of the fonts gave the best performance and the narrow Vialog performed poorly – he thus concludes that even though condensed fonts save space, it seems to be at the expense of legibility. This notion is confirmed by a distance study carried out by Garvey et al. (2001) showing a linear relationship with the most narrow faces delivering the worst performance and the widest ones the best performance. These findings are further supported by studies investigating width in the former standard fonts for American Highways, all suggesting that wide characters are more easily perceived at distance than narrow ones (Forbes et al., 1939; Zwahlen, 1995; Schnell, 1998) (fig.17).

While recognizing the general importance of serifs and hairlines\(^\text{18}\), Harry Carter emphasised that these features do not enhance the reading experience when used in type at distance with a few isolated words (H. Carter, 1931). This view was however not shared by David Kindersley, who found that typefaces for street signs, with heavy lines or even thickness, obstructs the open counters and lowers legibility when viewed from an acute angle (Dreyfus, 1957). Like others, Kindersley had observed that both corners and characteristic parts of letters have a tendency to round away and lose definition when viewed at a distance.

\(^{18}\) In most situations, the term hair-lines applies to the thin strokes of the Didone faces. However in this context, Harry Carter seems to be referring to serif typefaces of general contrast as opposed to Sans Serif typefaces of low contrast.
Kindersley’s solution, in the street signs and later in his proposal for British road and motorway sign system, was to apply serifs to the letters. He stated that ‘serif reinforces the individual character of the letter exactly where this loss is greatest’ (Kindersley, 1960, p.465).

Disputing the idea of emphasizing the corners in signage typefaces, Erik Spiekermann keeps the corners of the typeface Info rounded, arguing that round corners on back-lit signage typefaces make the shapes appear less distorted than if the angles were sharp (Spiekermann, 2006b). Gerard Unger supported such an approach in his 1974 back-lit signage typeface M.O.L. for the Amsterdam Metro (fig.18), based on the observation that when light shines through any opening of various shapes, it always tend to form a circle, and therefore, as in the case of Info, the corners of M.O.L. were rounded (Unger, 2007b).

Dwiggins invented his m-formula, Kindersley added serifs, Spiekermann and Unger rounded the corners: these designers were all trying to accommodate the loss of detail at distance yet dealing with it in very different ways.

While discussing corners in signage, a story relating to the origin of the Slab Serif typeface is interesting. Most historians seem to be somehow puzzled by the fact that the style was originally named Egyptian (Johnson, 1934; Tracy, 1994). According to Denman (1955), however, there might be a logical explanation for this. The story goes that during Napoleon’s Egyptian campaign, the army communicated by placing stations with intervals of a few miles. Their role was to paint messages on large boards that could be read by telescope from the next station that would then repeat the message for the following station, and so on. The letters used for this task were apparently Slab Serif faces, which due to the heavily squared serifs, appeared to be more distinguishable at a distance. However, the anecdote is contradicted by the fact that, at the time of the first Slab Serif and Sans Serif type specimens, the Slab Serif was not the only one being called Egyptian. For a while, the designations Egyptian and Antique were applied interchangeably between both Sans Serif and Slab Serif faces.
In the modernist movement of the 50s and 60s, the ideal typeface was a neutral typeface; this belief was mostly represented by Sans Serifs with closed apertures and a certain level of repetition of shapes. The signage typeface Airport by the company Crosby/Fletcher/Forbes – later renamed Pentagram – is an example of this (fig.19). As a redesign of the typeface Akzidenz the terminal endings were changed from diagonal to horizontal, resulting in rather closed counters since ‘this was considered an optical advantage for words used in large scale’ (Crosby/Fletcher/Forbes, 1970, p.16). However, most respected type designers today seem to agree that closed counters do not enhance visibility, especially when the font is viewed at a distance. This issue also concerned Adrian Frutiger, who stated that the horizontal ends of the terminals in Univers (fig.20) would have been different, were he to only design one bookface (Cheetham & Grimby, 1964). Frutiger explains the decision as being necessary, since the typeface was intended to expand and condense along a horizontal axis; however, when asked to apply Univers to the signage system for the Charles de Gaulle Airport, Frutiger could not see that it would work. Finding Akzidenz Grotesk, Helvetica and Univers unsuitable for the job, as their characters are not sufficiently differentiated, he therefore created a new type to meet his own demands for signage (Hunziker, 1998) (fig.21). The typeface was later reworked and extended to be released under his own name as ‘Frutiger’.

In 1912, the legibility researcher Barbara Elizabeth Roethlein set out to measure the legibility of a range of different fonts through distance threshold studies. In one of her experiments the focus was on isolated letters. In different studies the participants were exposed either to sheets showing letters in isolation, or were exposed to letters grouped in nonsense combinations.

Critics later argued that due to the single letters and the nonsense words, the test results did not provide any useful information regarding readability of cohesive reading (Tinker & Paterson, 1932; Ovink, 1938). Others targeted the use of the nonsense words, pointing out that the letter combinations were unpronounceable, and, therefore, that the words must have been spelled aloud, which by no means would simulate a normal reading situation (Pyke, 1926). These views were put forward before the theories of parallel letter recognition were known. At
that time, most legibility researchers supported the whole word model of word recognition since it was the best model given the data available. Another criticism of Roethlein’s study – one that is still voiced today in regards to legibility studies at distance – emphasized the fact that distance testing is very sensitive to the point size and heaviness of the font, and that this matter must have influenced the final test results (Starch, 1985), a factor that was acknowledged by Roethlein. Regardless of the criticism, Roethlein’s study does give us useful information on typeface weight and proportions. In a comparison of Cheltenham Old Style (Regular), Bold and Bold Condensed (fig.22), Roethlein found Cheltenham Bold to be legible at a longer distance than both the Regular and Bold Condensed fonts, a finding that suggests that bold weights give a better performance on distance than regular weights. Furthermore, a study of Roethlein’s list of The Average Legibility of Various Faces (fig.23), shows an almost even scale with the fonts of the largest x-height being most legible and the fonts of small x-height being less legible, indicating that characters of a large internal area and large x-height enhances the legibility in most fonts viewed at a distance. A further analysis of the findings suggests that bold weights and low stroke contrast additionally enhance distance visibility.

<table>
<thead>
<tr>
<th>Upper Case</th>
<th>Lower Case</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The Sixteen Roman Faces</strong></td>
<td></td>
</tr>
<tr>
<td>JENSON</td>
<td>281.7</td>
</tr>
<tr>
<td>BULFINCH</td>
<td>273.8</td>
</tr>
<tr>
<td>CHELT. W.</td>
<td>268.5</td>
</tr>
<tr>
<td>CENTURY O. S.</td>
<td>270.4</td>
</tr>
<tr>
<td>CLEARFACE</td>
<td>269.3</td>
</tr>
<tr>
<td>CHELT. O. S.</td>
<td>268.5</td>
</tr>
<tr>
<td>DELLA ROBBIA</td>
<td>260.8</td>
</tr>
<tr>
<td>NEWS GOTHIC</td>
<td>264.6</td>
</tr>
<tr>
<td>CENTURY EXP.</td>
<td>264.8</td>
</tr>
<tr>
<td>CASLON O. S.</td>
<td>250.7</td>
</tr>
<tr>
<td>CUSHING O. S.</td>
<td>247.6</td>
</tr>
<tr>
<td>DE VINNE NO. 2</td>
<td>243.2</td>
</tr>
<tr>
<td>RONALDSON</td>
<td>241.7</td>
</tr>
<tr>
<td>CUSHING MON.</td>
<td>228.4</td>
</tr>
<tr>
<td>CUSHING NO. 2</td>
<td>224.8</td>
</tr>
<tr>
<td>AMERICAN TYPEWR.</td>
<td>196.8</td>
</tr>
<tr>
<td>Average</td>
<td>282.8</td>
</tr>
</tbody>
</table>
The pixel

With type on-screen, especially in the small sizes, resolution and pixels play a central role. In the typeface Verdana, created for the screen by Matthew Carter, the terminals are cut off at a vertical angle to avoid the jagged pixels in the end-stroke of an angular ending, a feature that furthermore prevents closed counters (which would have been the result were the endings horizontal). Further, the typeface is generously fitted, has a high x-height and a low stroke contrast. While designing the Verdana/Tahoma/Nina and Georgia series, Matthew Carter created a bitmap version of particular target sizes first, and then afterwards designed the outline shape of the characters around the bitmaps (M. Carter, 2004b) (fig.24). Due to Microsoft’s ClearType technology, Carter did not have to do this in the newer Latin alphabet of the Meiryo typeface. The ClearType technology is based on a subpixel rendering system that controls the red, green, and blue (rgb) elements of each pixel (fig.25). In doing so, the aim is to optically enhance the resolution of the screen by controlling even smaller units within the pixel. Unfortunately, this effect only works on the vertical strokes; Microsoft therefore mixes subpixel rendering of the vertical strokes with an anti-aliasing of both vertical and horizontal strokes. Based on this mix of techniques, and to minimize the jagged diagonal in the between area of the characters, several of the designers behind the commissioned ClearType faces of 2004, enhanced the square feeling of the types, by emphasizing horizontal and vertical lines (Berry, 2004).

On print

In small print sizes and especially when printing is of bad quality, the shapes of the letter tend to melt on paper. At the time of manual punchcutting, each point size was cut individually, giving the gifted punchcutter the option of optically scaling the fonts to suit the requirements of the paper, the ink, and the human eye (fig.26). The adjustments made to the small sizes are of special interest in relation to legibility. However as pointed out by Harry Carter (1937), in the 16-18th centuries, 6- to 10 point types were cut solely for footnotes and for marginal notes accompanying a text set in a larger point size, and therefore the matter of creating small, highly legible fonts for longer passages of text was not an issue until the emergence of newspapers.
When optically compensating for small sizes, Walter Tracy (1986) suggested a general widening of the characters, moving the baseline a little lower on the body, which will result in shorter descenders and larger x-height characters. Harry Carter (1937) further advocated short ascenders, slightly heavier weight, low contrast, magnified strong serifs and an emphasis on the indentation at the junctions and the terminal of letters like ‘c’, ‘e’ and ‘a’. He stressed the importance of the white space inside the letters, which was his reason for recommending broadening the characters to preserve the balance between black and white.

One way of preventing the ink from dissolving the letterforms in the smaller sizes is by opening up the junctions in inktraps. An early example of this, where the outer side of the stroke is cut off at a straight angle, is demonstrated in the work of Johann Michael Fleischman (fig.27), a German punchcutter employed by Enschedé from 1743-68, a feature still present in several contemporary typefaces for small print, such as Jante by Poul Søgren and Parable (fig.28) by Christopher Burke.

The typeface Bell Centennial (fig.29) is famous for its inktraps. Created in 1978 by Matthew Carter and designed for the phone directories of AT&T, these inktraps differ from the ones applied by Fleischman, in having the edge of the strokes bending inwards instead of being cut off at the side. To enhance the visibility even more in the small sizes, Matthew Carter followed the advice of his father Harry Carter, mentioned above, and added square endings for emphasis to the terminals of the letters ‘a’, ‘c’, ‘e’, ‘s’, and ‘g’.

In 1993, when Martin Majoor tested a new typeface created for the Dutch telephone book, he established that the spikes and inktraps of Bell Centennial were no longer necessary due to the new printing techniques (Kinross 2002), a similar finding is further reported by Bruno Maag (2008). In contrast to this, Hoefler & Frere-Jones found that when designing the typeface Retina for the stock page of the Wall Street Journal, they still needed to create strong inktraps similar to the ones by Matthew Carter (Twemlow, 2004).
Different reading situations and different media influence legibility in ways that are not always the same. An opinion often expressed in the type community is that a typeface designed for small print is not intended for enlargement, and a typeface designed for signage is not to be applied as a text face, a belief supported by Hermann Zapf who emphasises that the purpose of a typeface determines its individual form, and it should therefore not be applied to other media (Zapf, 1987). Zapf’s own typeface Optima (fig.30) was initially designed for signage, but Zapf decided during the design process to change the purpose of the type to be a text face (Zapf, 1970). Nonetheless, regardless of Zapf’s intention, the typeface is today not only applied in text settings but is also often quite successfully applied to signage.

Erik Spiekermann argues against this notion of designing type for one purpose alone, pointing out that technical limitations, influencing the design of typefaces created for one purpose alone, can actually enhance the usability of the same type when applied under other difficult circumstances. He gives the example of his typeface Meta, designed for small sizes of print liable to smudge, and now also seen quite often on signs. He further points towards another of his typefaces, Officina, developed for office communication on laser printers, yet usable as a screen typeface as well (Spiekermann, 2006b) (fig.31).

It appears that the different qualities related to type applied on different platforms do not necessarily have to exclude one another. Another example is the typeface Gill Sans (fig.32) based on a sign of uppercase letters painted by Eric Gill on Douglas Cleverdon’s Bristol Bookshop, today being one of the most frequently applied Sans Serifs in running text.

It does seem possible to include features relating to several usages in one typeface, and in that way, to aim towards a more all-around design capable of covering several needs without seeming out of place. One must conclude this to be a positive outcome, since whatever the purpose a designer has in mind with his or her typeface, when out on the market, the chances are that the typeface will eventually appear in different media than those for which it was initially intended.
2.2

Empirical findings on differentiation of lowercase characters

Among earlier traditional legibility researchers, a popular subject was to study the relative legibility of letters by comparing the different characters within the alphabet.

Seven studies on the matter (carried out between 1885 and 1928) were later summarized by Tinker (1964): the methods employed were distance, short exposure, and the measurement of how far the letter could be recognised from the fixation-point of the eyes. Not all letters performed equally in the different test situations. Tinker found, however, the following rather consistent relationship within the seven studies:

Letters of high legibility: d m p q w
Letters of medium legibility: j r v x y
Letters of low legibility: c e i n l

While summarizing the work, Tinker points out that some distinguishing features of the letters aid legibility. Letters such as ‘b’, ‘d’, ‘p’, ‘q’ and ‘k’ all contain both a descending or ascending element and a well-defined x-height feature, making these letters among the most easily distinguishable.
All languages have their own individual word patterns and letter combinations. Pyke (1926) noted this fact when combining research material from 3 different nationalities (German, French and English), pointing out that some letters become more illegible when situated near certain other letters, and that the frequency of these letter combinations varies in different languages. The Cambridge Encyclopedia shows that in the English language the characters ‘e’, ‘t’, ‘a’, ‘i’, ‘n’, ‘o’, ‘s’, ‘l’, and ‘h’ are among the most frequently used (Crystal, 2003); as we shall see, this means that some of the most common characters also are among those most easily mistaken for each other.

Figure 34 lists a range of investigations focusing on the misreading of lowercase letters. Reviewing these studies, it appears that different typefaces are likely to have different letter misrecognitions. The study carried out by Bouma (1971) used the typeface Courier as the test material. Courier varies from other common typefaces by being monospaced and having very large serifs. This could be the reason for some of its misreadings not being common in other investigations: such as the ‘g’ being misrecognized as ‘q’, the ‘m’ being misrecognized as ‘n’, and the ‘w’ being misrecognized as ‘v’. The study by Geyer (1977) used the typeface Futura (fig.33) with its ‘t’ created out of a straight vertical stem and no curve at the bottom. This unusual shape probably increases the chances of the character being misread for ‘l’ and ‘i’.

However, a review of the data presented in figure 34 shows some pattern of recurring misreadings, with two main groups of troubled characters. One group is composed of the x-height characters of standard width built on a mixture of straight and curved lines (e-c-a-s-n-u-o), the other group is composed of the narrow letters with a single vertical stroke and a small width (i-j-l-t-f). These two letter groups will be the main subjects for the visibility investigation in the present project.

Based on the assumption that legibility is defined by the smallest amount of correlation in the surface of the characters, Legros & Grant (1912) set out to test the individual legibility in a range of different fonts, dealing with the most misread character pairs (fig.35). The work was carried out by calculating the number of units (one-thousandth of an inch) in a square covering each character in a font, and postulated that high legibility would be.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test methods</strong></td>
<td>Distance</td>
<td>Distance</td>
<td>Short exposure</td>
<td>Short exposure</td>
<td>Short exposure</td>
<td>Short exposure</td>
<td>Parafoval vision</td>
</tr>
<tr>
<td><strong>Typefaces</strong></td>
<td>Old Style</td>
<td>Courier</td>
<td>Didone Style</td>
<td>Old Style</td>
<td>Tactype</td>
<td>Courier</td>
<td>Old Style</td>
</tr>
<tr>
<td></td>
<td>Roman</td>
<td></td>
<td></td>
<td>Roman</td>
<td>Futura</td>
<td></td>
<td>Roman</td>
</tr>
</tbody>
</table>

* Highest frequency of mistaken at top  
** No specific order

<table>
<thead>
<tr>
<th>Misreadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>y &gt; p</td>
</tr>
<tr>
<td>i &gt; l</td>
</tr>
<tr>
<td>w &gt; v</td>
</tr>
<tr>
<td>h &gt; b</td>
</tr>
<tr>
<td>m &gt; w</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Misreadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>b &gt; h</td>
</tr>
<tr>
<td>p &gt; r</td>
</tr>
<tr>
<td>n &gt; a</td>
</tr>
<tr>
<td>h &gt; k</td>
</tr>
<tr>
<td>t &gt; i</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Misreadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>e &gt; c</td>
</tr>
<tr>
<td>l &gt; i</td>
</tr>
<tr>
<td>f &gt; r</td>
</tr>
<tr>
<td>k &gt; x</td>
</tr>
<tr>
<td>c + e</td>
</tr>
<tr>
<td>o + c</td>
</tr>
<tr>
<td>v &gt; r</td>
</tr>
<tr>
<td>q &gt; g</td>
</tr>
<tr>
<td>y &gt; r</td>
</tr>
<tr>
<td>j &gt; l</td>
</tr>
<tr>
<td>m &gt; u</td>
</tr>
<tr>
<td>c + o</td>
</tr>
</tbody>
</table>

Figure 34. Summarizing the work of five different researchers looking into misreading of lowercase letters.
equal to a low number of shared units across letters within the same font.

It is apparent that the repetition of shapes and low contrast, which dominate in the applied Sans Serif, performs rather badly under the circumstances elaborated above, whereas the more organic Blackfriars offers a much better result. Legros & Grant advocate the kind of serif applied in Old Style typefaces, arguing:

\[ \text{[...]} \text{that a heavy serif adds considerably to the non-coincident areas of the il, un, and bh pairs of lowercase characters (Legros & Grant, 1912, p.164).} \]

The Legros & Grant study is interesting in that it has a clear mathematical approach to the argument of enhancing the individuality of the characters. However, it fails to elaborate on the central issue of cohesiveness within the characters of a font. If a designer decided blindly to follow the Legros & Grant approach, he would obtain high scores in designing a typeface of characters with no characteristics in common; however this does not mean that the typeface would perform well in a normal reading situation, where some level of uniformity must be expected. Sanocki (1988) later demonstrated that reading performance is improved when the letters come from a single font instead of a mix of two fonts of different typefaces.

**Individuality of the x-height letters**

In a study of short exposure and distance threshold, the typographer scholar G.W. Ovink presented his participants with a
range of different styles of characters, and asked them to describe what shapes they saw, rather than to indicate the letters (fig.36). Ovink worked with only 6 participants. When applying a methodology that involves aspects of subjectivity, this low number of participants is somewhat questionable. However, in contrast to most other experimental legibility investigations, Ovink’s test material consisted of letter skeletons of similar weight, width and x-height.

Based on his study of three low contrast letter ‘s’ variations, Ovink concluded that the letter ‘s’ is the worst of the lowercase alphabet, since the features of the character themselves require exclusion from each other; his point being that when a diagonal movement is the aim, the curves become too small. However when the curves are the main element and even when the counters are large and open, the 8-shape appears resulting in possible misreadings. Yet one could argue that a misreading for the digit ‘8’ will never be an issue when the letter is part of a word. Ovink proposed a more angular shape with flattened curves; in serif typefaces he suggested that the serifs be kept at a minimum length.

In a study by Spencer and colleagues (1973a), the methodology applied was a gradually blurring of the images, which consisted of a range of different character variations of a low contrast Sans Serif font (fig.37). The study found that characters with extenders, in most cases, performed better than x-height characters, and that a one-storey ‘a’ with a tail, performed better than a two-storey ‘a’. This last finding contradicts the recommendations of both Ovink (1938) and Geyer (1977). While studying low contrast versions of the lowercase ‘a’, Ovink concluded
that the two-storey ‘a’ is more legible than the one-storey, and recommended a vertical movement in the round part in both styles. He further emphasised the importance of the loop in the two-storey ‘a’ not being too subordinate in narrow typefaces, mentioning the possibility of shortening the upper part, and in that way, maintaining emphasis on the loop.

The French researcher Emile Javal (1881) argued that when the top part of the ‘a’ has a round shape, it becomes easily mistaken for the lowercase ‘n’; to avoid this, Javal suggested a version, also found in the Italian Renaissance manuscripts, with a small top and a relatively long horizontal loop. To distinguish the characters ‘a’, ‘n’ and ‘u’ from each other, Sanford (1888) suggested keeping their openings at the top and bottom as wide as possible. This idea was followed up by Harris (1973), who tested Baskerville 169, Univers Medium and Gill Sans Medium (fig.38), and recommended open counters of ‘c’ and ‘e’ to avoid confusing them with ‘o’ and ‘a’. Like Harris, Tinker (1964) emphasised the role of the enclosed white space of a letter – arguing that the greater the size of the counter, the greater the legibility.

Tinker disagreed with the use of hairlines, his argument being that hairlines are most often applied to the essential parts that distinguish one character from the other, as in the crossbar of the ‘e’ causing it to be misread for ‘c’ (fig.39). However, Tinker further acknowledged that a thick crossbar, as seen in low contrast typefaces, could close up the eye of the ‘e’, and in that way cause poor legibility (Tinker 1928).

All three variations of the letter ‘c’ tested by Ovink had different sizes of apertures, the first being a circle cut along the radii of 195° and 135°, the second version supposedly a Futura ‘c’, and the last version having a small aperture and a squared serif at the top. Ovink found that the first and the last versions were immediately recognizable, but that the allegedly Futura ‘c’ with the open apertures was not as easily recognized, and was therefore concluded to be too narrow. This conclusion of Ovink contradicted the recommendation of Javal (1881), who stated that the ‘c’ should be similar to a half circle. In support of this, Tinker (1928) noted that the opening is the only characteristic

19) As pointed out by Christopher Burke (1998) this version of the ‘c’ is in reality just a circle cut off vertically and not the finally released version of the Futura ‘c’.

Figure 38. From the top, Univers Medium 689, Baskerville 169 and Gill Sans Medium. The fonts were applied in the study by Harris (1973, p.30).

Figure 39. Example of thin hairlines in the typeface Didot of 1991, designed by the Linotype staff and Adrian Frutiger, based on the work of Firmin Didot (1764-1836).
distinguishing ‘c’ from ‘o’, and consequently it should be as large as possible.

When discussing the letter ‘e’, Ovink (1938) emphasized the issues of avoiding restoration of the complete circle in the outside shape and keeping the counters open. This view was supported by several earlier researchers (Javal 1881; Sanford 1888; Legros & Grant 1916), who all advocated use of the oblique crossbar, making it possible to keep a large eye, while maintaining an open counter to the right (fig.40).

To differentiate the character ‘e’ from ‘o’, Javal (1881) proposed both an ‘e’ with a high crossbar near the top (fig.41), and an ‘e’ with a diagonal crossbar (fig.40). However, the recommendation of the high crossbar is contradicted by newer research carried out by Fox and colleagues (2007), finding that a lowercase ‘e’ with a high crossbar has a greater number of misreadings than a lowercase ‘e’ with a crossbar at the visual centre. This finding, in turn, was further supported by Harris (1973), who found the ‘e’ - ‘c’ confusion to be more likely in Baskerville with a high crossbar than in Gill Sans with a lower crossbar. Harris also found that serifs did not improve the legibility of certain characters, stating that:

> It appears that the serifs on the verticals of letters like h, n and u increases uncertainty about the identity of those letters. Confusions between b and h, and between n, u, o and a are significantly more likely in Baskerville than in the other faces (Harris 1973, p.32).

Javal (1881), Sanford (1888) and Tinker (1928) all recommended serifs to be short and preferably triangular, instead of linear in shape (fig.42). Sanford explained the function of the serifs as protecting the ends of strokes from the rounding effects of irradiation, and advocated the use of short serifs since they can achieve this outcome alone, while serifs that are too long easily result in letter confusion.

**Individuality of the narrow letters**

To create a room for the internal space, several researchers have stressed that width in characters seems to be preferable to narrowness (Javal, 1881; Sanford, 1888; Tinker, 1928). One should nevertheless be aware of the possibility of creating new forms
of misreadings when broadening normally narrow characters. This is evident in the study by Harris (1973) that found the misreading of ‘t’ for ‘c’ more common in the rather broad ‘t’ of Gill Sans than in the more narrow Univers and Baskerville versions. In general, however, both the letters ‘t’ and ‘f’ performed better in Gill Sans than in Univers, and Harris asserted this to be due to the slimmer versions of the characters in Univers, a view supported by Ovink (1938) who found both ‘f’ and ‘j’ to be more legible with a wide loop.

Both Javal (1881) and Ovink (1938) disapproved of the Didone style ‘t’ with the tail curling parallel to the stem (fig.43). To make the tail more prominent, Tinker suggested the same width of the stroke to the tail as to the stem (Tinker, 1928). Ovink (1938) recommended the height of the ‘t’ to be half way between the x-height and the ascending character height, and if the style of the typeface allows for it, then the ‘t’ should have a bracket on the left side of the crossbar to avoid a misreading of the ascending part for a dot. Both Javal and Ovink suggested the dot of the ‘i’ to be large in size and placed high above the stem. Ovink did not approve of serifs on the ‘i’, arguing that ‘serifs are superfluous, or even misleading if they are heavy’ (Ovink, 1938, p.33). Harris (1973), on the other hand, found in his experiment that the ‘i’ and ‘j’ of Baskerville were more legible than the two Sans Serif faces tested, and asserted this to be due to the serif emphasising the cap between the stem and the dot. The fact that the Baskerville typeface has a lower x-height, which places the dot further away from the stem, most likely played an additional role in the higher visibility of the Baskerville ‘i’ and ‘j’ in the Harris study.

A classic problem influencing many of the studies discussed is that when attempting to locate the most legible features in a comparison of typefaces of different proportions, weight, stroke, contrast, and look, the researcher is left with a number of variables influencing the results. With so many parameters varying in the test materials, it is difficult to identify the variables that influence the findings. The visibility investigation of the present project will avoid these difficulties.

---

Figure 43. The Didone style ‘t’ with a thin tail, as shown in the typeface Didot.
2.3

Designers and differentiation of lowercase characters

The empirical approach is motivated by the focus on one detail at a time. In contrast to this, a skilled type designer has the ability to view the work both as a whole and as a set of individual constituents. The saying often referred to by designers, that ‘type is a beautiful group of letters, not a group of beautiful letters’\(^{20}\), is an indication of how the overall feeling of a typeface is, in many ways, more important than the individual characters. Walter Tracy (1986) emphasised this in pointing out that since letters do not live in isolation, a focus on the regularity of the typeface’s texture is essential for holding comprehension over a longer period of time. The highly regarded Adobe designer Robert Slimbach, on the other hand, advocates some irregularity in asserting that ‘aesthetically pleasing variation can provide a mellowing effect that makes a type more comfortable to read’ (Adobe, 2007, p.5).

\(^{20}\) I have yet to locate the origin of the quotation. Gerard Unger refers to Matthew Carter as the source (Unger 2007a); however, when discussing the topic Carter himself notes that: ‘As the saying goes, type is a beautiful group of letters, not a group of beautiful letters’ (Cabarga 2004b, p. 200), indicating that he himself is not the creator. The quotation further appears as the title of an article by Steve Byers (Byers 2001), who lists a range of different typography related quotations, though with no reference to the origination of this specific one.
Throughout the history of letter design, the balance between regularity and individuality has been dealt with in various ways. In these pages, the main focus is on the enhancements of the characters’ distinctiveness, however, the overall harmony and balance of a typeface should never be forgotten.

Differentiation in the early days of printing

One of the earliest known sources showing a concern for the legibility of the Latin alphabet is by the English printer Joseph Moxon. In 1683 he expressed his interest by stating:

[…] we must conclude that the Roman letters were Originally invented and contrived to be made and consist of Circles, Arches of Circles, and straight Lines; and therefore those Letters that have these Figures, either entire, or else properly mixt, so as the Course and progress of the Pen may best admit, may deserve the name of true Shape, rather than those that have not (Moxon, 1683, vol. 2, p.15).

Moxon praised the Dutch-Letters of his time for their ‘commodious Fatness’, which, he argued, eases the reading process and ‘renders them more Legible’ (Moxon, 1683, vol. 2, p.15) (fig.44).

When describing the technical process of punchcutting, Moxon recommended that large letters should be counterpunched, while small letters should be engraved and sculpted. About 80 years later, the typefounder Pierre Simon Fournier described in his Manuel Typographique (published in 1764-68) how he used the same inner shape as counterpunch of characters ‘b’, ‘d’, ‘p’, ‘q’ and for ‘h’, ‘n’, ‘u’ (H. Carter, 1930) (fig.45). By reusing the counterpunch in several letters, the option of creating a differentiation between the counters of the characters was limited, so Fournier apparently approached the issue in another way, as he explained in his preliminary notice to the Modèles des Caractères of 1742:

I also have given the corners rather a squarer cut, and this I have done to some of the lower-case as well, and removed a certain roundness which was observable at the junction of vertical and horizontal strokes; this gives...
them an appearance of greater independence, separates the one from the other, and makes them more evidently distinct (H. Carter, 1930, p.289-290)²¹.

It appears, however, that not all printers were able to appreciate the subtle balance of distinctiveness and harmony that typefounders like Fournier struggled so much to create. In the 1755 editions of The Printer’s Grammar, the author John Smith described how printers living too far away from the founders for a regular supply, could replace certain letters for others if broken. Listing these letters, Smith explains how the characters b-q, p-d and n-u can be rotated and so replace each other; how ‘e’ could be changed to ‘c’ by cutting of the eye; how cutting off the ascender of ‘h’ would turn the character into an ‘n’ and that cutting the ‘n’ part off the ‘h’ would make it work as an ‘l’. At the time of writing, the Caslon Foundry was the most dominant in Britain, however, the serifs on the ascenders of ‘d’ and ‘b’ in these fonts, are quite different from the serifs on the descenders of ‘q’ and ‘p’, furthermore the ‘u’, unlike the ‘n’, has no serifs on the right site of the stems (fig.46). Based on this, Smith’s ideas were most likely not something William Caslon himself would have agreed upon. Probably for the same reason, in the 1787 edition of The Printer’s Grammar, the paragraph was removed.

In 1818, five years after the death of the typefounder and printer Giambattista Bodoni, his widow finished and published his Manuale Tipografico showing specimens of his type. Here, Bodoni discussed the regularity and harmony within the units of the letters, stating that:

[...] the standardisation of every thing which is not in itself distinctive, and the accentuation, so far as is possible, of the necessary marks of differentiation, will impart to all the letters a certain schematic regularity (Haddon Craftsmen, 1937, p.3).

Bodoni’s later work does not demonstrate a high concern for differentiation among letters, yet a look at his earlier fonts shows

2¹ A different translation is given by Geoffrey Dowding: ‘I squared the angles of these same capitals a little more, as well as some of the lower-case letters, where I removed a certain “rounding-off” in the angle between the perpendicular and horizontal strokes. This serves to give them more freedom, to distinguish one from another, and to make the strokes more clear’ (Dowding, 1961, p.63).
features that appear to confirm the statement above. In this example from 1788 (fig.47), the shoulder and the teardrop of the lowercase ‘r’ are rather heavy in the strokes compared to the similar parts of the ‘n’; the Transitional style difference in axis between the ‘o’ (vertical as in Didone faces) and ‘c’, ‘e’ (diagonal as in Old Style faces) further emphasises a distinction between the ‘o’ and the two other characters. However, the Didone style typefaces created by Bodoni and his contemporary Didot evolved over the years into fonts of high internal similarity between letters. It was a style that was generally accepted by the public, receiving relatively little critique (see more in Chapter 3.1). One such critique, however, was that of Citizen Sobre who, in 1800, explained his antipathy as follows:

The truth is that Garamond was careful to emphasise those parts of the shape of his types which distinguish them from one another—the ties for instance—while Didot emphasised those parts of the shapes of his types which are common to all (Morison, 1928, p.181).

Sobre further theorized that in the ‘u’ and ‘n’ of the design of Claude Garamond (fig.48) the focus is on the top part of the ‘n’ and bottom part of the ‘u’, and for these reasons ‘you cannot for a moment be in doubt as to which it is’. However, in the ‘u’ and ‘n’ of Didot’s fonts (fig.49). Sobre argued, the connecting parts are so thin in the hairlines that ‘you have to use discernment to avoid confusing’ (Morison, 1928, p.181).

The highlighting of parts of letters that are the same, and lack of emphasis on parts that separate one character from the other, seem to be present in both Bodoni’s and Didot’s later work. The difference between ‘e’ and ‘c’ in both designs relies solely on the hairline crossbar of the ‘e’, a fact further complicated by the teardrop of the ‘c’. The design of the ‘t’ is quite narrow in width, with an almost invisible, thin tail and a very short crossbar of rather similar proportion to the top serifs on the x-height characters. With the ascending part of the character being almost non-existent, it is not hard to imagine the letter being misread for ‘i’ or ‘r’. However due to the relatively low x-height, the dot of the ‘i’ is placed quite far from the stem – a feature that could emphasize its differentiation from ‘l’. 
Individuality of the x-height letters

A general change in the approach to typedesign came in the late nineteenth and early twentieth centuries with the Arts & Crafts movement’s rediscovery of the early Renaissance fonts. Designers were now no longer craftsmen following a certain movement in time, but creatives capable of working simultaneously within a range of different typeface categories, drawing inspiration from both historical and contemporary work, while adding their own ideas to the styles.

One of the twentieth centuries major designers, William Addison Dwiggins, advocated an open white area in the letter ‘a’. He further stressed that the generous openings within the characters ‘a’ and ‘e’ should be extended so that they have ‘clear white space both above and below their central strokes’ (Dwiggins, 1947, p.49). This was not the approach applied by early punchcutters such as Francesco Griffo (1450-1518) (fig.50) and Claude Garamond (1480-1561). In their day, the lowercase characters ‘a’ and ‘e’ had rather small eyes. Due to the quality of print and paper at that time, a small eye could easily be clotted with ink, so the Dutch founders who took over the leading role of type-founding after Garamond and his contemporary countrymen, opened up the counters slightly.

As mentioned in the previous chapter, applying a diagonal crossbar instead of a horizontal one to the lowercase ‘e’, makes the eye large while keeping the counter open, a feature inspired by the Humanist typefaces of Nicolaus Jenson (1420-80) (fig.51), favoured by Frederic W. Goudy (fig.52), and applied in the typeface Clearface (fig.53) by Morris Fuller Benton, designed for high legibility based on differentiation of characters. However, if the typeface is applied in running text, the diagonal crossbar will possibly work against the horizontal flow. Walter Tracy generally disliked the slanted bar of the ‘e’, especially in Sans Serif faces, arguing that it disturbs the stability of a word and creates an unwanted restlessness (Tracy, 1986).

In most typefaces, terminals on the lower parts of the letters ‘c’ and ‘e’ are similar in shape; however, in the typeface Futura (fig.54) by Paul Renner, the terminal is cut off horizontally in...
the 'e', to create a closed counter, and cut off vertically in the 'c', to create a fairly narrow shape. According to Burke (1998b), the decision was based on the German language having many 'ch' and 'ck' letter combinations; so by creating vertical stroke endings to the letter 'c', the characters could be tightly spaced and so minimize disrupting gaps in the word pattern.

As shown in Futura, the 'a' and 'g' of several German designs in the early twentieth century were created as one-storey characters. This adaptation of the one-storey style in Sans Serif faces can be attributed to the German population at the time being used to Gothic typefaces with one-storey characters, and more unaccustomed to the two-storey sort of the Latin alphabet. In addition to that, the simplicity of the one-storey characters is better suited for the low contrast of the Sans Serif faces (Monotype Recorder, 1927).

**Individuality of the narrow letters**

The dots on the 'i' and 'j' are in danger of connecting with the stem and being perceived as ascending characters, a risk more evident in Sans Serif than in Serif typefaces. Although not approved of by some designers - who argue that serifs do not belong on a Sans Serif typeface - several newer Sans Serif typefaces have a slab serif on 'i' and 'j', a feature originating in monospaced typewriter faces, where the narrow characters must cover the same area as the broader ones. This phenomenon is today most present in typefaces designed for electronic media and signage, where optimal character recognition is thought to be essential and where spacing is less subtle (fig.55).

This slab serif on 'i' and 'j' is often found in Sans Serif faces by Erik Spiekermann, who further shows a liking for the Sans Serif 'l' with a tail to the right. In the typeface created by Edward Johnston for the London public transport system in 1916, the curly tail on lowercase 'l' is quite dominant (fig.56). Johnston and his collaborators at the London Electric Railway Company aimed at producing a new type with letters of high individuality (Howes, 2000). As Johnston some years before had explained in his influential book *Writing & Illuminating & Lettering*:

---

22) According to Tracy, a third of the typeface listed as 'Neudeutsche Schriften' in the 1926 volume of the Handbuch der Schriftarten have single-storey a's and g's (Tracy 1986).
The “Characteristic Parts” are those parts which most particularly serve to distinguish one letter from another. We should therefore, when constructing letters, give special attention to their preservation, and sometimes they may even be accentuated with advantage (Johnston, 1906, p.247-250).

Johnston applied a historically inspired calligraphic tail to the lowercase ‘l’ (fig.57), a somewhat uncommon feature in typodesign at the time. Walter Tracy later recognized this intention of differentiating the lowercase ‘l’ from uppercase ‘I’, however he emphasised that when placed in words, the character ‘was so broad that the letter stood aloof from the one that followed’ (Tracy, 1986 p.89). When, in 1974, Tracy was asked to redesign Johnston Sans he condensed the broad ‘l’ noticeably.

Another way of enhancing the ‘l’ is shown in the Romain du Roi typeface developed between 1693-1745. At the x-height, and to the left of the stem, the lowercase ‘l’ was given a spur quite similar to the serif on the x-height characters, the feature can also be seen in lowercase ‘l’ in textura gothic (fig.59), and is present in some French calligraphy of the 17th century (fig.60). Hermann Zapf (1987) speculates that the idea behind might be to distinguish the character from the uppercase ‘I’.

In the first years of printing, a common adaptation from the calligraphic hand was to place the dot of the ‘i’ slightly ahead of the stem; the attribute was present in the work of both the masters Jenson and Griffo and later, although less radically, in the work of Garamond. Except for a subtle presence in Jenson revivals such as Centaur and Adobe Jenson, the attribute is not commonly applied in contemporary designs (fig.61).

For a harmonious spacing of characters, the long Old Style terminals on ‘I’ and ‘j’ requires kerning with neighbouring characters (fig.62). In metal and wood type, a kerned letter would have a terminal that projects outside the body, making it heavily exposed to damage and therefore not very popular among
founders\textsuperscript{24} (fig.63). Due to the kerning difficulties and the high number of various ‘f’ ligatures that compositors had to work with, in 1805, Charles Earl Stanhope presented a list of possible improvements to ease the work with stereotyping, among these being a proposal for an alteration in the terminal of the letter ‘f’ (Hart, 1896). The idea was to create a kernless ‘f’ by curling the terminal heavily inwards in a way where the right site of the stem bends backwards to make room for the narrow counter - a look that was already somewhat present in the work of Baskerville, enhanced by the Didone typeface designers, fulfilled by Lord Stanhope\textsuperscript{25}, and taken to the extremes in the nineteenth century fat faces. It was for practical reasons, rather than to improve legibility, that the ‘f’ got to be narrow (fig.64).

The coming of mechanical typesetting meant that the need for kernless characters became even more evident, since kerning in the Linotype line-caster machine was impossible. The challenge for designers creating type for the line-caster was therefore to find a way in which the kernless characters would harmonize properly with the rest of the typeface. The typeface Sabon (fig.65) created by Jan Tschichold in 1967\textsuperscript{26} is an example of this. The typeface was an interpretation of the work by Garamond, which meant that the broad Garamond ‘f’ had to be narrowed to accommodate the lack of kerning facilities. When, in 2002, Jean François Porchez created the updated version, Sabon Next (fig.66), he went back to the original wide lowercase ‘f’, believing that that would have been the intention of Tschichold, were he to create the typeface with today’s technical possibilities of controlling kerning and fitting of the letters (Berry, 2006, p.35). Tschichold highly admired Garamond’s

\textsuperscript{24} Kerned letters being attended with more trouble than other Sorts, Founders are sometimes sparing in casting them; whereas they rather require a larger number than their Casting-Bill specifies, considering the chance which Kerned letters stand, to have their Beaks broke, especially the Roman f, when it stands at the end of a line, where it is exposed to other accidents, besides those from the lie-brush’ (Smith, 1755 p.34).

\textsuperscript{25} Lord Stanhope, on the reception of his new letter shape: ‘Man is so much the child of custom, and so much the implicit admirer of fancied beauty, that I believe if the human body generally was very round-shouldered, and if the head projected considerably beyond the chest, it would, in such case, be deemed a deformity to see a man with an upright body, and carrying his head erect. Having this opinion upon so weighty a subject, I was not surprised to meet with objectors to the proposed alternation in the shape of so humble a servant of literature as the letter f: readers had been so long accustomed to meet her with a downcast head, apparently too weighty to be supported by her feeble neck, that she failed in meeting with a welcome reception in assuming the appearance of strength by carrying her head upright’ (Hansard 1825, p. 477).

\textsuperscript{26} The typeface was commissioned to work both for hand-compositing and for Linotype and Monotype composing machines.
original design (fig. 67). He praised the tail of Garamond’s ‘t’ for enhancing legibility, and emphasised that the top of the ‘t’ ‘is right in appearing to be short’ (Tschichold, 1969, p. 53). He further criticised contemporary faces for lengthening the ascending part of the ‘t’, arguing that it enhances the possibility of a misreading for the lowercase ‘l’.

In both Slab Serif and Sans Serif faces, the lowercase ‘f’ has a tradition of being rather narrow in width. This is partly due to their origination in sign writing, where space tends to be more limited in width than in height, and partly due to their time of arrival in the printing industry corresponding with the already existing fashion of a slim ‘f’ in Serif faces. The emphasis on the narrowness of already narrow characters is evident in Paul Renner’s Futura. Not only does the typeface have a remarkably slim ‘f’, but also the characters ‘j’ and ‘t’ have no tails at all. The concern that these unusual letterforms would result in incoherent word patterns was raised by several of Renner’s contemporaries, but Renner himself was convinced that the constructed geometric base that the typeface was built upon would, as he stated, connect the ‘many individual marks into unity of forms’ (Burke, 1998b, p. 98).

Discussing individual letterforms within a typeface is a difficult task since changing one feature of a design often influences a range of other elements. As Frederic Goudy (1940) stressed, not all letters that seem legible in isolation work when forming words. Different neighbouring characters can influence the performance of a letter in different ways. Factors like these make the study of the individual letter isolated from the word a somewhat complicated matter. A designer will always have to work on the whole and on the individual elements simultaneously. In that sense, the empirical approach and the practical approach to typeface legibility appear to contradict each other in ways that can be highly problematical. It would be rather difficult to successfully apply, in a slavish manner, all empirical results to one typedesign and still keep it harmonious. Creating a well designed, legible typeface is therefore not about creating dissimilar characters; the aim must be to find, amongst other things, an optimal balance between uniformity and differentiation and, through that, to attain legibility.
Three new fonts

Three new fonts were developed as the test materials used in the visibility investigation of the present project. The fonts TinkerTest, OvinkTest and PykeTest are all named after twentieth century legibility researchers. Each of the fonts has one main function. OvinkTest is to be applied on signage, PykeTest is to be applied in running text, and TinkerTest is designed for shorter paragraphs. That being said, type designers often have no knowledge of what context their work will appear in, in the future. As pointed out by Gerard Unger, since you rarely have a clear picture of the reader, ‘There is not a lot you can do to meet the wishes of individual groups of users’ (Unger, 2007a, p.110). The intention with the three finished typeface families is therefore to create new fonts that not only function when applied in their main situations but also have the ability to adapt to other usages if necessary.

The three fonts vary on several levels. One such difference lies in the internal proportion between the ascenders, descenders and x-height characters (fig.68). The signage font OvinkTest has a large x-height. PykeTest on the other hand, is inspired by Giambattista Bodoni’s work, giving it a low x-height emphasising the extenders. The three fonts further vary in the strokes

Figure 68. The typeface PykeTest is about 43%, TinkerTest about 52% and Ovink about 59% of the total letter height.

27) For theories on writing tools: see Noordzij (2000, 2005) and Johnston (1906).
with TinkerTest being influenced by the brush; PykeTest being influenced by the pointed nibbed pen, while OvinkTest is more constructed in the shapes. OvinkTest is furthermore a Sans Serif face; PykeTest a Serif face, and TinkerTest is a hybrid between Serif and Sans Serif styles.

Being a hybrid, the possibilities of varying the use of serifs are higher in TinkerTest than in the two others. As shown by Harris (1973), serifs on the lowercase counters appear to lower visibility. In TinkerTest, serifs are therefore removed from the counters. To further enhance the differentiation of 'b' and 'h', the right stem of characters 'h', 'n', 'm' extend in an arc to the right, which ideally should enhance the horizontal flow of the font as well. The focus on differentiation of characters also applies to the uppercase alphabet of TinkerTest. To avoid the mis-reading of 'D' as 'O', the stroke of the bowl extends to the left of the stem in the upper part of the ‘D’, for harmony the feature is also applied to other uppercase characters with a horizontal
upper stroke (BEFPR). The possible problem of misreadings between 'R' and 'B' is accommodated for by an arc attached to the right of the leg of the 'R' – the arc is further applied to the other similar shaped characters: 'A', 'K' and 'k'.

Inspired by the street signage font of the Dane, Knud V. Engelhardt (fig.72), the main attribute of the Sans Serif OvinkTest is the similarity of forms. OvinkTest is a Sans Serif font strongly based on repetition of shapes; the challenge here is to keep the feeling of uniformity while still differentiating the individual characters on a subtle level.

PykeTest is a text face that combines Bodoni inspired features with a horizontal movement in the characters (fig.73). To improve readability and minimize the vertical feeling, the ascending lowercase stems are bent slightly forward, the thin

Figure 72. Lettering for the street signage of Gentofte Kommune by Knud V. Engelhardt, 1923.
hairlines and the radical transition between horizontal and vertical counters have been toned down, and the lowercase letters ‘e’ and ‘c’ have a diagonal axis.

--

A readable typeface is often described as a Humanist Sans Serif or Old Style serif face – both based on the broad nibbed pen stroke with a medium level of contrast. A reference to this tool in the letterforms will automatically lead to some form of differentiation between characters. Although TinkerTest has reminiscence of the broad nibbed pen, none of the three fonts in this study are based completely on that. The decision not to directly include a traditional Humanistic typeface was motivated by a challenge to see whether legibility can be reached by other means than the standard Humanistic style. Through these new fonts, I wish to celebrate the diversity in typography by creating typefaces that are not based on the most obvious shapes; the task, then, is to make them legible while maintaining their own style and personality.

Figure 73. A Bodoni type specimen of 1788 with a reminiscence of Old Style features.
2.5 The study of character shapes in three new fonts

To control for visibility in the later familiarity investigation, and to enhance the legibility of new typefaces, I studied how to improve the design of individual letters. The three new fonts created for the test material each contain several variations of the most frequently misrecognized letters.

In contrast to other legibility studies where characters from different fonts are compared with each other, the characters in this investigation are compared with alternate variations within the same font. This controls not only obvious variables such as size, proportion, contrast, and weight, but also the look and personality of the typeface – all matters that often vary in font-to-font studies. Controlling these parameters enhances the probability of locating the exact features that improve visibility. By empirically testing the three new fonts under the same conditions, each with their own range of variations of the selected letters, the intention is to discover whether the different look of the font can influence the results, or whether there are more universal conclusions to be made. This methodology has the advantage of being able to better understand the changes that result in any performance differences, but it does also have the disad
vantage of not being able to examine the full variety of designs that are seen in the real world. A concern with looking only at variations within a single typeface is also that the conclusions may only apply to that typeface. However, by empirically testing three rather different fonts under the same conditions, we can be more certain that the findings will broadly apply to letter recognition.

As earlier discussed (see Chapter 1.4), any legibility investigation will benefit from several test methods. The two methods chosen here are both threshold studies; one method is based on recognition at distance; the other is a short exposure method focusing on parafoveal vision. By applying these approaches, issues related to both signage and text typefaces are under investigation.

We know that both foveal and parafoveal visions are important to continuous reading, yet the short exposure methodology applied in the present study did not detect any errors of identification when test material was placed in the foveal, and so the focus was on the parafoveal alone.

**Test material**

In the study of the two main groups of troubled letters (e-c-a-s-n-u-o and i-j-l-t-f), two categories of letter variations were created: those that are quite similar to familiar letterforms, and those that might appear unusual to the casual observer.

**Familiar letter variations**

The goal of all the letter variations is to create greater distinction between letters. With a few exceptions, similar letter skeletons were tested on each of the three fonts. The variations of the letter ‘i’ in the TinkerTest and OvinkTest faces focus on different levels of serifs. The serifs emphasize the separation of the stem from the dot, and are expected to have better visibility than the versions without serifs. Serif faces need serifs on the ‘i’, therefore there was no reason to test these variations on PykeTest. For similar reasons the tailless ‘u’ was not tested in PykeTest because it is aesthetically too out of place in a Serif face.

A high level of differentiation between ‘n’ and ‘u’ is expected to improve visibility. To study this hypothesis, u2 has no tail and the bowl of versions n2 and u3 detaches closer to the middle of
the stem than does versions \( n_1 \) and \( u_1 \). In doing so it is expected that focus will be directed towards the areas where the letters are most different from each other. A similar diagonal stroke is represented in the crossbar of version \( e_2 \). This is expected to improve the recognition rate by opening up the counter.

The Law of Closure described by the German school of Gestalt psychologists, suggests that our perceptual system tends to complete incomplete shapes by filling out caps. Following this hypothesis it would be expected that the smaller the aperture in ‘c’ and ‘e’ the larger the risk that the eye will close the cap and mistake these letters for ‘o’. The hypnosis is further studied in open and closed apertures of the letter ‘s’ in OvinkTest. Following the same idea, the familiar two-storey ‘a’ has versions with open apertures, and versions with more closed apertures. It would be expected that closed apertures result in terminals optically joining the bowl and then lower legibility.

The one-storey ‘a’ was tested in TinkerTest and OvinkTest. Due to the dominating x-height round shape, this version would be expected to show a low level of legibility and a high number of misreadings for the lowercase ‘o’.

The narrow letters (l-f-t-j-i) cover a small horizontal area. It would be expected that if spread over a larger area the visibility of the characters will improve. Letters of this group all have wide and narrow versions tested.
Figure 75. The TinkerTest letter variations.

Figure 76. The OvinkTest letter variations.

Figure 77. The PykeTest letter variations.
Unfamiliar letter variations
Some of the tested variations were more unusual than others; these more unfamiliar versions can be divided into two main groups. One approach explores the possibility of extending the height of the character; the other the possibility of adding uppercase character shapes to the lowercase alphabet.

Many lowercase letters use neither the ascending or descending space. The approach investigates the inclusion of the ascending and descending areas of some of the letters that do not usually make use of this space. Normally being x-height characters, the ‘a’ and ‘s’ move above and below this area. We know that larger sizes are more easily perceived than smaller sizes at a distance, so by extending the ‘a’ into the ascending area and the ‘s’ into the descending area, it would be expected that the otherwise highly compact inner spaces of the characters open up and become more distinctive.

During the evolution of the lowercase alphabet, the early uncial pen hands mixed present day upper and lowercase alphabets. Inspired by this tradition, the letter variations n3, e5 and t3 are uppercase letters reduced to x-height characters – the hypothesis goes that these already recognized letterforms could replace the existing lowercase versions, and still function in combination with the uppercase alphabet.

Study 1

Methods
The first study applied a methodology of short time exposure of a single character.

Participants
There were a total of 41 participants in this study. Not all participants saw all three fonts. 15 only saw TinkerTest, 2 only saw OvinkTest, 18 saw OvinkTest and PykeTest, 3 saw TinkerTest and OvinkTest, and 3 saw TinkerTest and PykeTest. TinkerTest and PykeTest were each exposed to 21 participants, where OvinkTest was exposed to 23 participants. Most of the participants were compensated with a gratuity of Microsoft software or hardware. Some early participants received no compensation.
The participants included 26 students with art and design backgrounds from the Royal College of Art, and 15 students from the Imperial College. Their ages ranged from 19 to 34 with an average age of 25.7 years. The participants came from a variety of backgrounds (British, French, Brazilian, Danish, Canadian, Swedish, Norwegian, Spanish, Slovenian, Polish), and all self-reported either normal or corrected-to-normal visual acuity. Because the mean number of errors made by participants from the two schools was not reliably different, the data from the two groups will be reported combined.

**Material**

The test material was created in Macromedia Flash MX and shown on a 15-inch MacBook Pro laptop with a screen resolution of 1440 x 900 pixels set to maximum brightness. The three fonts (TinkerTest, OvinkTest and PykeTest) were all presented with anti-aliasing at the vertical size of 45 pixels (an Em-square of about 1 cm)\(^{28}\). To minimize eyestrain caused by the background light of the screen, the background colour was a shaded white (#E6E6DD) with the presented letters in black (#000000). The ambient room light was typical for an office environment.

**Procedures**

The parafoveal area is important for continuous reading (Rayner, 1978; Rayner, McConkie & Ehrlich, 1978). Characters that are easily identified in the parafoveal area result in higher levels of legibility in running text. Test materials were therefore located 2 cm to the right of the fixation point where participants placed their focus. Their eyes were placed at a distance of 50 cm from the screen.

Each character variation within a font was presented 3 times per participant. To maintain an approximately equal appearance between the 26 letters of the alphabet, the 15 characters of the English alphabet that were not under investigation were each exposed 5 times – all occurring in the same random order.

The instruction was to focus on a red dot on the screen and then press the space key to trigger an exposure of a single character, which participants were asked to name. Each letter was

---

\(^{28}\) Since this is not a study of comparison between typefaces, the three faces are not adjusted according to x-height.
exposed for a period of about 43 milliseconds. To ease the participant into the test, a selection of the characters not under investigation were presented as the first 5 exposures. A mask (exposed for 43 milliseconds) of randomly placed black dots followed directly after each letter exposure: this removes the afterimage on the retina and controls the timeframe in which the image in reality would appear on the retina. Participants were informed that they would be presented uppercase and lowercase letters, they were not asked to hurry their response, as their responses were not timed.

**Results of short exposure study**

Each letter variation for TinkerTest and PykeTest was presented a total of 63 times, and for OvinkTest 69 times. If the participant correctly identified the presented letter, the trial was counted as correct. A chi-square test was applied to the raw totals of correct identifications compared with errors, i.e. a two-dimensional table. Tests were only conducted between variants within a font, as it was not a goal of this investigation to compare differences between the font.

*For visualization of the complete data set, see Appendix A.*

<table>
<thead>
<tr>
<th>Letter ‘a’</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(3)=39.8, p=.0001$</td>
<td>$\chi^2(3)=73.87, p=.0001$</td>
<td>$\chi^2(2)=1.99, p&gt;.05$</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Versions</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>a1</td>
<td>8</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>a2</td>
<td>13</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>a3</td>
<td>36 *</td>
<td>57 * †</td>
<td>-</td>
</tr>
<tr>
<td>a4</td>
<td>11</td>
<td>16</td>
<td>24</td>
</tr>
</tbody>
</table>

* Post-hoc tests showed reliably more errors than each of the other versions
† A high frequency of misreadings for the letter ‘q’ (20)
### Letter ‘c’

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(2)=3.04, p=.05$</td>
<td>$\chi^2(2)=1.09, p=.05$</td>
<td>$\chi^2(2)=6.59, p=.05$</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

### Versions

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>5</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>c2</td>
<td>8</td>
<td>14</td>
<td>19</td>
</tr>
<tr>
<td>c3</td>
<td>3</td>
<td>18</td>
<td>16</td>
</tr>
</tbody>
</table>

### Letter ‘e’

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(4)=47.94, p=.0001$</td>
<td>$\chi^2(4)=37.82, p=.0001$</td>
<td>$\chi^2(4)=26.43, p=.0001$</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

### Versions

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>e1</td>
<td>6</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>e2</td>
<td>8</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>e3</td>
<td>12</td>
<td>10</td>
<td>17</td>
</tr>
<tr>
<td>e4</td>
<td>13</td>
<td>23 **</td>
<td>25 ***</td>
</tr>
<tr>
<td>e5</td>
<td>35 * †</td>
<td>32 **</td>
<td>36 **</td>
</tr>
</tbody>
</table>

* Post-hoc tests showed reliably more errors than each of the other versions
** Post-hoc tests showed reliably more errors than versions 1, 2, 3
*** Post-hoc tests showed reliably more errors than version 2
† A high frequency of misreadings for the letter ‘c’ (21)

### Letter ‘n’

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(2)=0.29, p=.05$</td>
<td>$\chi^2(2)=0.81, p=.05$</td>
<td>$\chi^2(2)=1.09, p=.05$</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

### Versions

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>n1</td>
<td>9</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>n2</td>
<td>7</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>n3</td>
<td>8</td>
<td>10</td>
<td>21</td>
</tr>
</tbody>
</table>

---

**Figure 80.** Results and raw number of errors in study 1 for the letter ‘c’.

**Figure 81.** Results and raw number of errors in study 1 for the letter ‘e’.

**Figure 82.** Results and raw number of errors in study 1 for the letter ‘n’.
Letter 's'

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(1)=1.37$, $p&gt;.05$</td>
<td>$\chi^2(2)=7.61$, $p=.02$</td>
<td>$\chi^2(1)=5.21$, $p=.02$</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Versions</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>s1</td>
<td>9</td>
<td>13 *</td>
<td>9 **</td>
</tr>
<tr>
<td>s2</td>
<td>-</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>s3</td>
<td>4</td>
<td>15 *</td>
<td>2</td>
</tr>
</tbody>
</table>

* Post-hoc tests showed reliably more errors than version 2
** Post-hoc tests showed reliably more errors than version 3

Figure 83. Results and raw number of errors in study 1 for the letter 's'.

Letter 'u'

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(1)=0.18$, $p&gt;.05$</td>
<td>$\chi^2(1)=0.05$, $p&gt;.05$</td>
<td>$\chi^2(1)=0.53$, $p&gt;.05$</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Versions</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>u1</td>
<td>2</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>u2</td>
<td>4</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>u3</td>
<td>-</td>
<td>-</td>
<td>12</td>
</tr>
</tbody>
</table>

Figure 84. Results and raw number of errors in study 1 for the letter 'u'.

Letter 'f'

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(2)=5.12$, $p&gt;.05$</td>
<td>$\chi^2(2)=1.49$, $p&gt;.05$</td>
<td>$\chi^2(2)=4.28$, $p&gt;.05$</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Versions</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>f1</td>
<td>14</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>f2</td>
<td>6</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>f3</td>
<td>15</td>
<td>23</td>
<td>7</td>
</tr>
</tbody>
</table>

Figure 85. Results and raw number of errors in study 1 for the letter 'f'.
### Letter ‘i’

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(2)=2.52$, $p=.05$</td>
<td>$\chi^2(2)=12.98$, $p=.002$</td>
<td>-</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>no</td>
<td>yes</td>
<td>-</td>
</tr>
</tbody>
</table>

#### Versions

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>i1</td>
<td>36</td>
<td>39</td>
<td>-</td>
</tr>
<tr>
<td>i2</td>
<td>35</td>
<td>49 *</td>
<td>-</td>
</tr>
<tr>
<td>i3</td>
<td>43</td>
<td>28</td>
<td>-</td>
</tr>
</tbody>
</table>

* Post-hoc tests showed reliably more errors than version 3

### Letter ‘j’

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(1)=4.51$, $p=.03$</td>
<td>$\chi^2(1)=4.79$, $p=.03$</td>
<td>$\chi^2(1)=1.59$, $p&gt;.05$</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

#### Versions

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>j1</td>
<td>10</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>j2</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

### Letter ‘l’

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(1)=1.15$, $p&gt;.05$</td>
<td>$\chi^2(2)=0.75$, $p&gt;.05$</td>
<td>$\chi^2(1)=6.46$, $p=.01$</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

#### Versions

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>l1</td>
<td>37 †</td>
<td>42</td>
<td>45</td>
</tr>
<tr>
<td>l2</td>
<td>30</td>
<td>37</td>
<td>30</td>
</tr>
<tr>
<td>l3</td>
<td>-</td>
<td>40</td>
<td>-</td>
</tr>
</tbody>
</table>

† A high frequency of misreadings for the letter ‘t’ (25).

---

**Figure 86.** Results and raw number of errors in study 1 for the letter ‘i’.

**Figure 87.** Results and raw number of errors in study 1 for the letter ‘j’.

**Figure 88.** Results and raw number of errors in study 1 for the letter ‘l’.
Table 8. Results and raw number of errors in study 1 for the letter ‘t’.

<table>
<thead>
<tr>
<th>Chi-square</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2(2)=62$, $p&gt;.05$</td>
<td>$\chi^2(2)=2.92$, $p&gt;.05$</td>
<td>$\chi^2(2)=0.57$, $p&gt;.05$</td>
<td></td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

Study 2

Methods

The second study used a distance threshold methodology to study the legibility of the same font letter variants. Our expectation was that the findings would be roughly similar to the findings in the first study.

Participants

There were 41 participants in study 2, though 7 were disqualified because they did not meet the minimum visual acuity requirement of being able to recognize stimuli at a distance of 4.5 meters. This left 34 participants. All three fonts were each exposed to 20 participants. The participants were compensated with a gratuity of Microsoft software or hardware.

Material

The fonts, computer, and environment were identical in the two studies.

Procedures

In this investigation, the laptop was placed on a podium at the eye level height of a standing person of about 175 cm. The angle of the screen was adjusted to fit the given height for each person.

The first presented character was the letter ‘d’. As identified by Tinker (1928) this character is one of the most easily recognized letters. The purpose of this first exposure is to locate the individual vision threshold. The participant was placed at a distance of 10 meters from the screen, and asked to move slowly forward until the presented letter was barely identifiable; this
was the distance – varying from 4.5-9 meters (with an average of 6 meters) from the screen – at which the individual participant was tested. From this distance, participants were asked to name each of the letter stimuli. A new letter was presented on screen after each participant response. Participants were not asked to hurry, and were permitted to take as many breaks as they felt necessary.

This methodology is different from the one applied in other recent distance studies, such as those by Sheedy and colleagues (2005) and the studies of the Clearview typefaces (Garvey, Pietrucha & Meeker, 1997), where the maximum distance is measured for each letter, and the distance itself becomes the data rather than the accuracy from a particular distance. However, the Sheedy and Clearview methodology does not identify which other letters the character tested is most likely to be misread for – a parameter that is measurable with the present methodology.

**Results of the distance study**

Each letter variation was presented a total of 60 times. If the participant correctly identified the presented letter, the trial was counted as correct. The inferential statistics of a chi-square distribution were conducted on the raw totals of correct observations.

*For visualization of the complete data set, see Appendix A.*

<table>
<thead>
<tr>
<th>Letter ‘a’</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>aaad</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>aaa</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>aaa</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>aaa</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Versions</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>a1</td>
<td>19</td>
<td>19</td>
<td>27</td>
</tr>
<tr>
<td>a2</td>
<td>33 **</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>a3</td>
<td>49 * †</td>
<td>48 * †</td>
<td>-</td>
</tr>
<tr>
<td>a4</td>
<td>19</td>
<td>12</td>
<td>16</td>
</tr>
</tbody>
</table>

* Post-hoc tests showed reliably more errors than each of the other versions
** Post-hoc tests showed reliably more errors than versions 1 and 4
† A high frequency of misreadings for the letter ‘o’. TinkerTest a3 (23), OvinkTest a3 (22).

![Figure 90. Results and raw number of errors in study 2 for the letter 'a'.](image)
### Letter 'c'

<table>
<thead>
<tr>
<th>Versions</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>19</td>
<td>13</td>
<td>23 *</td>
</tr>
<tr>
<td>c2</td>
<td>9</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>c3</td>
<td>13</td>
<td>13</td>
<td>30 *</td>
</tr>
</tbody>
</table>

* Post-hoc tests showed reliably more errors than versions 2

<table>
<thead>
<tr>
<th>Chi-square</th>
<th>$\chi^2(2)=5.82, p&gt;.05$</th>
<th>$\chi^2(2)=5.87, p&gt;.05$</th>
<th>$\chi^2(2)=39.10, p=.0001$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistically</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

### Letter 'e'

<table>
<thead>
<tr>
<th>Versions</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>e1</td>
<td>45</td>
<td>29</td>
<td>38</td>
</tr>
<tr>
<td>e2</td>
<td>40 †</td>
<td>34</td>
<td>48 †</td>
</tr>
<tr>
<td>e3</td>
<td>39</td>
<td>30</td>
<td>42 †</td>
</tr>
<tr>
<td>e4</td>
<td>44</td>
<td>44 ** ††</td>
<td>44</td>
</tr>
<tr>
<td>e5</td>
<td>49</td>
<td>40 †</td>
<td>55 *</td>
</tr>
</tbody>
</table>

* Post-hoc tests showed reliably more errors than versions 1, 3 and 4
** Post-hoc tests showed reliably more errors than versions 1 and 3
† A high frequency of misreadings for the letter 'c'. TinkerTest e2 (21), OvinkTest e5 (23), PykeTest e2 (20), Pyketest e3 (21).
†† A high frequency of misreadings for the letter 'o' (40).

<table>
<thead>
<tr>
<th>Chi-square</th>
<th>$\chi^2(4)=5.43, p&gt;.05$</th>
<th>$\chi^2(4)=11.52, p=.02$</th>
<th>$\chi^2(4)=15.13, p=.004$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistically</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

### Letter 'n'

<table>
<thead>
<tr>
<th>Versions</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>n1</td>
<td>9</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>n2</td>
<td>13</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>n3</td>
<td>16</td>
<td>17 *</td>
<td>36 *</td>
</tr>
</tbody>
</table>

* Post-hoc tests showed reliably more errors than versions 2

<table>
<thead>
<tr>
<th>Chi-square</th>
<th>$\chi^2(2)=2.47, p&gt;.05$</th>
<th>$\chi^2(2)=6.46, p=.04$</th>
<th>$\chi^2(2)=36.81, p=.0001$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistically</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

### Figure 91. Results and raw number of errors in study 2 for the letter 'c'.

### Figure 92. Results and raw number of errors in study 2 for the letter 'e'.

### Figure 93. Results and raw number of errors in study 2 for the letter 'n'.

[94] TYPEFACE LEGIBILITY: TOWARDS DEFINING FAMILIARITY
Letter ‘s’

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(1)=12.09, p=.0005$</td>
<td>$\chi^2(1)=10.05, p=.007$</td>
<td>$\chi^2(1)=51.39, p=.0001$</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

Versions

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>s1</td>
<td>38 **</td>
<td>22</td>
<td>46 **</td>
</tr>
<tr>
<td>s2</td>
<td>-</td>
<td>11</td>
<td>-</td>
</tr>
<tr>
<td>s3</td>
<td>18</td>
<td>27 *</td>
<td>8</td>
</tr>
</tbody>
</table>

* Post-hoc tests showed reliably more errors than versions 2
** Post-hoc tests showed reliably more errors than version 3

Figure 94. Results and raw number of errors in study 2 for the letter ‘s’.

Letter ‘u’

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(1)=0.17, p&gt;.05$</td>
<td>$\chi^2(1)=4.51, p=.03$</td>
<td>$\chi^2(1)=0.16, p&gt;.05$</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

Versions

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>u1</td>
<td>15</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>u2</td>
<td>18</td>
<td>16 +</td>
<td>-</td>
</tr>
<tr>
<td>u3</td>
<td>-</td>
<td>-</td>
<td>20</td>
</tr>
</tbody>
</table>

Figure 95. Results and raw number of errors in study 2 for the letter ‘u’.

Letter ‘f’

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(2)=4.11, p&gt;.05$</td>
<td>$\chi^2(2)=0.58, p&gt;.05$</td>
<td>$\chi^2(2)=21.54, p=.0001$</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

Versions

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>f1</td>
<td>32</td>
<td>30</td>
<td>28 *</td>
</tr>
<tr>
<td>f2</td>
<td>21</td>
<td>29</td>
<td>28 *</td>
</tr>
<tr>
<td>f3</td>
<td>26</td>
<td>26</td>
<td>7</td>
</tr>
</tbody>
</table>

* Post-hoc tests showed reliably more errors than version 3

Figure 96. Results and raw number of errors in study 2 for the letter ‘f’.
**Letter ‘i’**

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(2)=9.68$, p=.008</td>
<td>$\chi^2(2)=5.66$, p&gt;.05</td>
<td>-</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>yes</td>
<td>no</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Versions</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>i1</td>
<td>15</td>
<td>16</td>
<td>-</td>
</tr>
<tr>
<td>i2</td>
<td>29 *</td>
<td>27</td>
<td>-</td>
</tr>
<tr>
<td>i3</td>
<td>30 *</td>
<td>27</td>
<td>-</td>
</tr>
</tbody>
</table>

* Post-hoc tests showed reliably more errors than version 1

**Figure 97.** Results and raw number of errors in study 2 for the letter ‘i’.

**Letter ‘j’**

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(1)=24.08$, p=.0001</td>
<td>$\chi^2(1)=27.27$, p=.0001</td>
<td>$\chi^2(1)=2.33$, p&gt;.05</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Versions</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>j1</td>
<td>29</td>
<td>31</td>
<td>4</td>
</tr>
<tr>
<td>j2</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

**Figure 98.** Results and raw number of errors in study 2 for the letter ‘j’.

**Letter ‘l’**

<table>
<thead>
<tr>
<th></th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\chi^2(1)=19.21$, p=.0001</td>
<td>$\chi^2(2)=6.30$, p=.04</td>
<td>$\chi^2(1)=0.00$, p&gt;.05</td>
</tr>
<tr>
<td>Statistically reliable</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Versions</th>
<th>TinkerTest</th>
<th>OvinkTest</th>
<th>PykeTest</th>
</tr>
</thead>
<tbody>
<tr>
<td>l1</td>
<td>43 †</td>
<td>38 †</td>
<td>41</td>
</tr>
<tr>
<td>l2</td>
<td>18</td>
<td>33</td>
<td>40 ††</td>
</tr>
<tr>
<td>l3</td>
<td>-</td>
<td>46 *</td>
<td>-</td>
</tr>
</tbody>
</table>

* Post-hoc tests showed reliably more errors than versions 2
† A high frequency of misreadings for the letter ‘i’. TinkerTest l1 (30), OvinkTest l1 (24), OvinkTest l3 (31).
†† A high frequency of misreadings for the letter ‘t’ (20).
Discussion

This technique of comparing letter variations within a typeface has provided insights about letter legibility that were not previously available. The majority of the studies presented in Chapter 2.2 examine legibility by comparing typefaces; these studies struggle to make comparisons because every letter differs in several dimensions. The present investigation is complementary to this work by investigating letters that come from the same typeface with many fewer differences, which allows us to be more confident in understanding why one letter performs better than another.

The x-height letters

As expected, the performance of the one-storey a₃ [α α] was generally bad, with recurrent misreadings for letters ‘q’ and ‘o’. So does this mean that a one-storey ‘a’ should never be used? In relation to the inexperienced reader it does appear to have a purpose. Recognition is a dominant factor when learning to read, the fact that the one-storey ‘a’ references to the letter shape that most children learn to write, has a positive influence on the inexperienced reader (Sassoon 2001). The present study, however, focuses on the experienced reader, where references to one’s own writing hand are less essential.
The two-storey ‘a’ versions did not show a reliable difference in performance, except for the TinkerTest distance study, where a2 [a] performed reliably poorer than versions a1 [a] and a4 [a]. This is an unexpected difference. A possible reason for this might originate in the shape of the bowl. The upper part of the bowl in version a2 [a] is more diagonal in TinkerTest than in the two other fonts [a a], it furthermore bends slightly inwards, disrupting the dynamic movement of the curve, and making it look more like a spine than a bowl.

The idea that detaching the bowl from the stem would enhance visibility of versions n2 [n n n] (and PykeTest version u3 [u]) has not been confirmed – showing no statistically reliable difference over n1 [n n n] in any situation. The tailless version u2 [u u] gave a poor performance in OvinkTest at distance, however in other situations it presented no statistically reliable difference from u1 [u u].

The initial hypothesis that closed apertures in ‘c’ and ‘e’ would lower visibility was not confirmed in the case of the letter ‘c’ – only showing a statistically reliable difference between the open c2 [c] and the more closed versions c1 [c] and c3 [c] in the distance study of PykeTest. The PykeTest c1 and c3 are the only versions tested with a teardrop on top, this finding suggests that teardrops do not improve visibility at distance. It further appears that in the parafoveal vision, when the letter ‘c’ is viewed in isolation, the viewer registers the cut off area in the circle regardless of the size of the area and therefore, in contrast to all existing recommendations, showed no difference in the characters having closed or open apertures. This matter also seems to have influenced the performance of versions s1 [s] and s2 [s] of the OvinkTest, showing an advantage in favor of the closed apertures of s2 [s] in both the short exposure and the distance studies. The fact that the OvinkTest s1 [s] has a diagonal spine and s2 [s] a rounded spine might be the reason for the advantage towards version s2 [s]. It appears that the shape of the spine actually had a larger influence on the visibility of the ‘s’ than the apertures being opened or closed, a finding that, contradicts the recommendations of the scholar G.W. Ovink (1938), who suggested a diagonal spine of the ‘s’. The surprising performances of TinkerTest a2 [a] and OvinkTest s2 [s] might therefore be related. It seems that a diagonal stroke in the bowl and spine of these letters lowers their visibility and that these areas would benefit from being more rounded in shape.
The hypothesis that closed apertures of the letter ‘e’ lower visibility was confirmed in both OvinkTest and PykeTest at short exposure and in OvinkTest at distance. The three remaining familiar ‘e’ variations showed no internal differences. Yet all versions expect e1 demonstrate a high number of misreadings for the letters ‘c’ and ‘o’.

The narrow letters
Based on the findings it seems reasonable to recommend wide versions of narrow letters. All distance studies presented a small difference between t1 [t t t] and t2 [t t t] in favor of the latter, although only with a statistically reliable difference in TinkerTest. The t2 [t] of TinkerTest is exceptionally wide. As Harris (1973) has pointed out, broadening the ‘t’ seems to eliminate the misreading with other narrow characters; it does, however, also create a new dominant misreading for the letter ‘c’. Compared to a narrow version, it appears that a reasonable wide ‘t’ is more visible at a distance.

The broad j2 [j j j] delivered a good performance on all accounts in TinkerTest and OvinkTest; however, no statistically reliable difference was demonstrated between j1 [j] and j2 [jj] in PykeTest. The broad ‘j’ is particularly successful because it does not introduce any new confusions. The broad version l2 [l l l] also showed a reliably better performance in the Tinker distance study when compared to the narrower l1 [l], and in the OvinkTest distance study compared to the straight stem l3 [l]. These are the hypothesized results. More surprising, however, are the results of the PykeTest short exposure study, showing a reliably better performance with the curved tailed l2 [l] than the more common serif style l1 [l].

Applying the broad variations of j2 [j j j] and l2 [l l l] in a typeface will possibly result in spacing problems: j2 [j j j] will overlap with descending characters to the left, an issue causing potential trouble in the Scandinavian languages having a high number of gj letter combinations. Version l2 [l l l] would create a disrupting area of extra white space when placed to the left of another stem. When implement these variations in a final typeface, it could be necessary to apply a number of extra ligatures and kerning pairs.

The descending f3 [f f f] showed no difference in TinkerTest and OvinkTest, however a reliably better performance was dem-
onstrated in PykeTest at distance. This result may be due to the f3 [f] version of PykeTest being broader in shape than the f3 [f f] version of the two other fonts. Contrary to the broad versions of the ‘j’ and ‘l’ groups, the wide f2 [f f f] did not perform reliably better than any of the other tested variations.

The hypothesis that serifs on the letter ‘i’ improve visibility was confirmed. In both the TinkerTest and OvinkTest distance study, i1 [i i] with the slab serif on top, was recognized more often than i2 [i i] and i3 [i i]: however, only with TinkerTest showing a statistically reliable difference, it seems as if the slab serif on top of the stem helps to clarify the letterforms, although when placed at the bottom, the character becomes too wide. This appears, however, not to have been the case in the OvinkTest short exposure study, where i3 [i i] performed reliable better than i2 [i i].

The unfamiliar letter shapes
The reader’s expectation with the exposed character does seem to influence the more unfamiliar versions tested. Taking that into account, the hypothesis predicting the inclusion of the extending areas to improve visibility of the ‘a’ and ‘s’ was confirmed. As a consequence, the extending versions a4 [ã ã a] and s3 [$$ s $$] delivered a rather good performance. In the case of TinkerTest and PykeTest distance studies, and the PykeTest exposure study, s3 [$$ s $$] showed a reliably better performance than the x-height s1 [s s], and a4 [ã ã a] showed in general no statistically reliable differences compared to other two-storey ‘a’ versions.

When uppercase shapes were applied as x-height lowercase letters (versions e5, n3, t3), the results were mixed. The e5 [E E E] variation performed poorly. It appears that the upper and lower crossbars are over dominating the middle crossbar, which in some cases resulted in a high number of misreadings for the letter ‘c’. Version n3 [N N N], on the other hand, showed no noticeable difference in most of the studies, except for PykeTest at distance presenting a statistically reliably bad performance compared to both versions n1 [n] and n2 [n], and in OvinkTest at distance favouring version n2 [n].

In all three fonts, the version t3 [T T T] was frequently misread for the letter ‘r’ in the distance threshold study, and delivered a statistically reliably poorer performance compared to other versions of the ‘t’. On the other hand, in the short exposure study, this kind of misreading was non-existent, and the three versions of the letter performed in general quite similarly.
Implementing the high performing unfamiliar versions in a font within a new typeface would theoretically place the font on an equal visibility level to a font of familiar letterforms within the same typeface. The two fonts will have the same level of visibility but very different familiarity levels. Studying readers’ experience with these different versions will be the subject of the familiarity investigation presented later in this project.

---

There are many differences between a letter from one typeface and the same letter in another typeface. In contrast to the font-to-font methodology, the present methodology of studying within-font matters provides data that has a practical use for the design of new typefaces.

Very few of the findings in this project appear to be unanimous between all fonts and test methods – a fact that presents a textbook example of how difficult it is to make universal statements on legibility related matters. Each font shows findings that relates only to that specific font, and each test method shows results that are different than those of the other method. This confirms the notion that some aspects are important to distance viewing and others are important to the parafoveal vision. Chapter 2.1, however, suggests that although some visibility related issues are directly related to the usage and the application of the type, others can be transferred between platforms.

Trying to summarize the findings of the present study, it appears that visibility in typefaces for distance viewing will be improved if a serif is placed on the top of the stem of the letter ‘i’; that the middle part of the ‘s’ and the two-storey ‘a’ in general benefits from being more rounded than diagonal; that the unfamiliar ascending and descending versions of the same letters are of equal visibility as the familiar letterforms; that a one-storey ‘a’ has low visibility; that the x-height uppercase ‘n’ and the x-height uppercase ‘t’ in the parafoveal view are of equal visible as the lowercase versions; that the crossbar of the ‘e’ should be placed at the visual centre with an open aperture; that a closed aperture of the ‘c’ is equally as legible as an open aperture in the parafoveal view; that a teardrop on ‘c’ lower the visibility at distance; and except for the lowercase ‘f’, narrow characters benefit from being slightly broadened.
2.6

Implementing results in three new typefaces

The test material in the visibility investigation was intended to help study letter skeletons with as few interfering variables as possible. What happened was that other shapes, originally not assumed to influence the results, actually ended up doing so. Consequently, when finishing each typeface, I decided not only to implement the findings that were related to that specific typeface but also to adapt results from the testing of the two other fonts. The round spine of the OvinkTest s2 [S] is an example of that, performing reliably better than the more diagonal spine of OvinkTest s1 [S]. Based on this finding, it seems reasonable also to curve the spines of the ‘s’ in the typefaces Tinker and Pyke. Another example is the wide version of f3 in PykeTest, suggesting that the commonly applied descending ‘f’ in Italic faces might benefit in general from being slightly widened.

The visibility investigation further demonstrated that narrow characters gain from being broadened, however the study was solely concerned with single letters, and did not look into the effect different letterforms have on words. As pointed out by Walter Tracy (1986), an extraordinarily wide lowercase ‘l’ could have a negative influence on the important word pattern. How-
ever, one must assume that this only goes for situations in which another letter is placed to the right of the 'l', and not when the 'l' is the last letter of a word. The lowercase 'l' of the typeface Tinker is therefore narrow when a character follows directly after, and broader when it is followed by a space (fig.101).

As discussed in Chapter 2.1, the early type founders adjusted the fonts to suit the point sizes. Following this tradition, two additional versions of the typeface Pyke were created: one for small text and one for display sizes. Compared to the original version developed for regular text sizes of 9-14 points, the version for small sizes (PykeMicro) has a larger x-height, smaller contrast and broader letters (fig.102). The version for display sizes has a smaller x-height, larger contrast and narrower letters.

Historical traditions dictate Didone style narrow characters to be extra narrow. Nonetheless, to follow the findings of the visibility investigation and to enhance the horizontal flow, this tradition was disregarded in the text versions of Pyke; instead, emphasis was given to the forward movement by broadening and opening the loops in both characters 'j' and 'f'. Another disregarded Didone style feature in the text and micro versions was the heavy teardrops on the letters 'f', 'j', 'a', 'c' and 'r', which seem to enhance the vertical movement in the letters. The situation is different in PykeDisplay. Since headlines and titles in printed matters often appear in large sizes at a close reading distance, legibility may in these situations be less significant than in versions for running text. Furthermore, headings rarely consist of longer paragraphs, and are often perceived in few fixations. A horizontal emphasis is consequently not as vital, and the Didone style tradition of the narrow letters 'j' and 'f' can, in PykeDisplay, be applied without troubling the reader, and so can the heavy teardrops. These decisions have resulted in the text and micro versions moving towards the Transitional typeface category, while the Display versions are more Didone style fonts.

The original OvinkTest font was developed to be economic in space when applied on signs; however as discussed in Chapter 2.1, research suggests that fonts of generally broad letters can be recognized from a larger distance than fonts of more narrow width, and that the weight of a font is also an essential factor in perception at distance. Based on these findings, the typeface

Figure 101. The lowercase ‘l’ in TinkerRegularRoman.

Figure 102. In 44 point size from the top: PykeMicro, PykeText and PykeDisplay, all in the fonts RegularRoman.
family Ovink does not include Light or Condensed fonts. Instead, it has an extra range of Expanded fonts that are slightly wider than the OvinkNormal fonts, which are to be applied when horizontal space is not a factor.

The primary function of the typeface Tinker is for short paragraphs; the secondary function is a more all-round use. The finished family includes a variety of weights ranging from ExtraLight to Bold, some of the more extreme of these weights will not function on all platforms. At a distance, the ExtraLight weights will vanish, and in small print sizes the Bold weights will melt and dissolve.

An objective in the development of the three typefaces was to try not to eliminate typeface categories and weights according to what is most legible. Some categories have a lower level of incorporated legibility than others, and some weights can only be applied in specific situations. Instead of eliminating these more problematic versions, my aim has been to provide fonts of different typeface categories with the highest possible level of legibility – and then leave it to the graphic designer to implement the typefaces wisely in the final layout.
Figure 103: The finished version of the typeface family Tinker contains 10 fonts, in 5 weights from ExtraLight to Bold, in Roman, Italic, Small Caps, and lining and Old Style figures. All weights further have an OpenType stylistic setting featuring unfamiliar letter variations, and are equipped with a Central European Character set.

For the full type specimen see Appendix B.
Figure 104: The finished version of the typeface family Ovink contains 16 fonts, in regular and bold weights in Positive and Negative, and in Normal and Expanded width. All of these are in Roman, Italic, and lining and Old Style figures. All fonts further have an OpenType stylistic setting featuring unfamiliar letter variations, and are equipped with a Central European Character set. For the full type specimen see Appendix C.
<table>
<thead>
<tr>
<th>Font Style</th>
<th>Sample Characters</th>
</tr>
</thead>
<tbody>
<tr>
<td>PykeMicroRegularRoman</td>
<td>ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz</td>
</tr>
<tr>
<td>PykeMicroRegularItalic</td>
<td>ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz</td>
</tr>
<tr>
<td>PykeMicroBoldRoman</td>
<td>ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz</td>
</tr>
<tr>
<td>PykeMicroBoldItalic</td>
<td>ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz</td>
</tr>
<tr>
<td>PykeTextRegularRoman</td>
<td>ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz</td>
</tr>
<tr>
<td>PykeTextRegularItalic</td>
<td>ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz</td>
</tr>
<tr>
<td>PykeTextBoldRoman</td>
<td>ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz</td>
</tr>
<tr>
<td>PykeTextBoldItalic</td>
<td>ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz</td>
</tr>
<tr>
<td>PykeDisplayRegularRoman</td>
<td>ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz</td>
</tr>
<tr>
<td>PykeDisplayRegularItalic</td>
<td>ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz</td>
</tr>
<tr>
<td>PykeDisplayBoldRoman</td>
<td>ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz</td>
</tr>
<tr>
<td>PykeDisplayBoldItalic</td>
<td>ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz</td>
</tr>
</tbody>
</table>
Section 3

FAMILIARITY
3.1

The long-term influence of familiarity

Since the creation of the first book, letterforms and typeface styles have moved in a variety of directions, controlled by different movements and by individuals with a will to exploit or standardize lettering in various ways. An analysis of some of the public responses to these developments, and their relation to time and place, would help clarify the kind of influences that changes in typography can have on the present day reader.

Writing in 1919, William Addison Dwiggins (1947) stated that a reader should never be conscious that the page is made up of letters, and that if any single character presents itself to the reader as a single character, it will interrupt the flow of the reading. He later goes on to stress that:

"A type that stops you in the middle of a sentence and asks you to admire its smartness is a bad type (Dwiggins, 1928, p.17)."

The traditionalist Stanley Morison expressed a similar view. Known for his strong beliefs, Morison announced that:
any disposition of printing material which, whatever the intention, has the effect of coming between author and reader is wrong (Morison, 1930, p.61).

Morison further believed in maintaining the already known letterforms, stating that for typefaces to be satisfactory the essential forms have to ‘corresponds with that handed down’ (Morison, 1924, p.59).

In the influential essay Printing Should be Invisible, by Morison’s colleague Beatrice Warde, Warde compared typography with a crystal goblet. With the wine symbolising the message of the text and the goblet symbolising the typography, she argued that the goblet should always be as clear as possible since ‘no cloud must come between your eyes and the fiery heart of the liquid’ (Warde, 1936, p.6). Warde elsewhere argued that in order to forget that a type is there we must be familiar with it (Warde, 1927). In his later more conservative years, Jan Tschichold supported this opinion. Criticising characters such as the first Futura ‘r’ for making the reader trip up while reading (fig.106), he explained:

Our type is in fact an absolutely inflexible form, and offers no possibility for any but minute alternations if we still want to read with ease (Tschichold, 1969, p.53).

Describing the effect of familiarity, Gerard Unger (2007a) points to the fact that when we always do things the same way, we no longer have to think about the action. In regards to reading, this makes it much easier to focus on the content and switch to the unconscious automatic form that is an essential element of a successful reading experiment. This notion is supported by the ‘Predictive Coding’ hypothesis put forward by neuroscientists. The idea is that instead of passively registering visual inputs, the brain actively makes a prediction based on prior experiences, and by that, anticipates the visual input (Hosoya et al., 2005; Gjedde, 2008). This phenomenon makes it possible to carry out

![Figure 106. Drawings described as Paul Renner’s first designs for Futura (Burk, 1998, p.87).](image)
a task without being conscious of it, a mode that can be interrupted by an unexpected event, such as missing a step on the staircase, or stumbling over an unusual character in the reading of an otherwise predictable typeface.

As we now shall see, the available historical reference material on typeface innovation suggests that the level of exposure and the level of common letterforms, among other factors plays a significant role in influencing the long-term typeface familiarity.

Uncommon letterforms
Before the western world settled on the standard Latin alphabet, the style of the scripts differed noticeably between each other. These scripts not only varied in their overall appearance but were occasionally also built out of rather different letterforms.

Reading and writing in the Middle Ages
Reading skills in the early Middle Ages were obtained by the selected few having some form of connection with cloisters, monasteries and castles, and were not a part of the life of the general public. According to Denman (1955) an ordinary copy of the Bible would cost a common labourer the equivalent of 15 years’ salary. The lack of reading and writing skills among ordinary people is further emphasised by the fact that a song based on a defeat of Charlemagne in 778 was first written down in the later part of the eleventh century (Anderson, 1992). Reading and writing were not altogether easy tasks: a study of manuscripts from the Middle Ages shows that the same word tended to be spelled in different ways by different scribes, who wrote down in a phonetic manner what they were hearing (Anderson, 1992). Like the scribes, the reader was part of a culture dominated by verbal communication; a medieval reader usually gained new knowledge orally and was not familiar with the processing of information from paper. In the course of understanding the meaning of a text, references were therefore more connected to whether one had heard the word before than to whether one had seen the word written down before (Chaytor, 1945). This practice was not made easier by the fact that many manuscripts were hastily written and full of contractions that
did not particularly enhance the word recognition process. The norm among ancient readers was to read the text out loud, as is shown in a note left by an eighth century scribe:

Three fingers hold the pen, the eyes see the words, the tongue pronounces them as they are written and the body is cramped with learning over the desk (Chaytor, 1945, p.14).

Not being able to read a text in silence is a lack of competence we today associate with people of poor reading skills. If this was the case – which Chaytor documents in a number of references – an average reading session would have been of a rather troubled nature. The ancient reader would not have been able to look past the book hand and allow the lettering to become transparent in the way most legible typedesigns do today; instead the reader would be far more aware of the actual act of reading, a matter resulting in a performance of quite a tiresome nature.

After the fall of the Roman Empire, national scripts flourished all over Europe, leading Charlemagne, in 789, to develop and standardise the Carolingian Minuscules (fig.108). In this process, some readers had to adjust to a more unfamiliar script style. That Charlemagne succeeded in standardising a script for his empire is possibly more related to the poor reading skills of his people than to any other factor. If you already feel uncomfortable with the script you normally read, the shifting to another script is not the biggest interruption in your life. The fact that the tradition was to apply different calligraphic hands side by side with texts for different purposes further shows that the writing style was never transparent; the typographic presentation was just as essential as the content itself. So due to the low number of readers, and that poor reading skills seem to have been the norm, legibility in this period appears to have been of a somewhat secondary value.

30) In the late eight century the Emperor Charlemagne decided to appoint the Anglo-Saxon Benedictine monk Alcuin to apply one script as a standard in all manuscripts; the script decided upon was based on existing written hands, and was to be known as the Carolingian Minuscule – the style that later became the Roman lowercase alphabet.
Gothic and Latin types

The Italian rediscovery of the Carolingian Minuscule in the Renaissance did not immediately result in the rest of Europe automatically following the change from Gothic to Latin. However, the spirit of humanism did eventually penetrate all areas within Western Europe, and so did the Roman and Italic scripts – a process that apparently was not without difficulties. For instance, the Dutch punchcutter Joos Lambrecht ran into problems in 1539 when attempting to introduce Roman types to the public, noting:

I am ashamed about the uncivilized attitude of so many people in our country, who are unable to read our low-Dutch or Flemish tongue when printed in Roman type, saying that they do not recognize the letters, and that it seems Latin or Greek to them (Middendorp, 2004, p.17).31

This view was also found in Germany many years later. In a letter written in 1882, Otto von Bismarck stated that it took him much longer to read a page set in Latin type than if the same page was set in Gothic (Burke, 1998b).

The Gothic typefaces dominated the German speaking part of Europe up until the Second World War. In 1904, Rudolf von Larisch published a pamphlet criticising the lack of differentiation and the over-complexity of the Gothic typefaces (Kinross 2004): a view supported by German scholars who found that the dominant use of Gothic created problems for them when working in international collaborations. Due to the Gothic style being unfamiliar to intellectuals of other nationalities, German scholars went in opposition to the national norms and printed their books and periodicals in Latin typefaces instead (Steinberg, 1979).

Nonetheless, when the National-Socialist Party came into power, the freedom of typeface choice was no longer an option; the permitted fonts were only Gothic typefaces32. However in

31) A slightly different translation of the same text is presented by Goldschmidt: ‘I am ashamed of the clumsiness that in our country so many people are found who cannot read our Netherlands Dutch or Flemish when printed in Roman letters...’ (Goldschmidt, 1950, p. 45).

32) The reason for this choice seems to be based on the fact that central and northern European countries, at this time, were the only places to see Gothic in print, so the National-Socialist Party used this notion to promote the Gothic alphabet as a pure Nordic symbol (Lehmann-Haupt, 1954).
1941, the Nazis changed their view by declaring, that 'the so-called Gothic script consists of Schwabacher-jewish letters' and therefore the 'Fuehrer has decided that Roman type from now on shall be designated as the normal type' (Lehmann-Haupt, 1954, p.172).

The reason for this sudden u-turn is most likely related to difficulties in communicating Nazi propaganda in the defeated nations, and in that case being forced to prioritise the reader’s familiarity with the type over ideological concerns. A notion supported by a letter circulated 10 days later and written by the National-Socialist Party, stating that all newspapers and periodicals were to be printed only in Roman ‘because foreigners who can read the German language, can hardly read this script’ (Burke, 1998b, p.165). Although Gothic typefaces are still to be found on beer labels, album covers and traditional store signage in German speaking countries, due to the association with the Nazis, the typefaces have not been widely used since the war.

According to the Swiss, Adrian Frutiger - who generally expresses scepticism towards the whole legibility issue - the reading of Gothic type makes no difference as long as the reader is used to the style. Frutiger emphasises that he himself had no difficulties in learning to read with Gothic characters (Eurographic Press Interview, 1962). The idea that those familiar with reading the Gothic typefaces can easily perceive the characters, is further supported by the notion that three legibility studies carried out by different German researchers around 1920, all found a small advantage towards the Gothic style when compared to Latin style typefaces (Tinker, 1923). It does not seem that the extreme similarity in the vertical strokes of the characters was a serious issue for readers at that time. As long as they had experience with the style, according to Tinker, they actually preferred the monotony of the Gothic type to the more differentiated forms of the Roman faces.

Common letterforms

With a few exceptions, the skeleton of the letters in the Latin alphabet has stayed constant for more than 500 years. Since the Renaissance printers however, a range of new typeface styles have come to life.
John Baskerville

The story of John Baskerville is an often-cited historical reference when arguing for the power of the familiarity effect. Zužana Licko, who designed the typeface Mrs. Eaves (fig.109) as a free interpretation of Baskerville’s type, notes that:

When selecting a typeface for revival, I recalled reading in various sources that Baskerville’s work was severely criticized by his peers and critics throughout his lifetime and after. From personal experience, I could sympathize (Licko, 2001, p.68).

Given that Baskerville’s ink was darker in colour, the paper brighter and the type sharper than what was produced by his contemporaries, the argument goes that since people were not accustomed to the look, they disapproved of a typeface that today is viewed as being among the most legible ever produced (fig.110).

It is true that Baskerville’s work did get mixed reviews, however it is unlikely that he was disliked quite as much as some writers have suggested. One of Baskerville’s contemporary admirers was the famous French punchcutter Pierre Simon Fournier who complimented Baskerville’s type for being ‘cut with much spirit’ and constituting ‘real masterpieces’ (Updike, 1937, vol.2, p.108).

Baskerville himself was highly aware of this recognition, writing in a preface to his second published book that:

After having spent many years, and not a little of my fortune in my endeavours to advance this art; I must own it gives me great Satisfaction, to find that my Edition of Virgil has been so favourably received (Baskerville, 1758, p.A3).

On several occasions over the years Baskerville emphasised in letters that he had obtained ‘the reputation of excelling in the most useful art known to mankind’ (Straus & Dent, 1907, p.100 & p.102).

Most of the criticisms of Baskerville came from his compatriots. In a book that presented all English foundries, written by one of Baskerville’s contemporaries, he gets the following brief mention:
Mr Baskerville of Birmingham that enterprising place, made some attempts at letter-cutting, but desisted and with good reason. [...] indeed he can hardly claim a place amongst letter-cutters. his typographical excellence lay more in trim glossy paper to dim the sight (Mores, 1778, p.86).

Another anecdote often referred to in connection with Baskerville’s type is based on a letter that Benjamin Franklin – also a printer – wrote to Baskerville in 1760. In the letter Franklin describes a conversation he had with a Gentleman:

[...] he said you would be Means of blinding all the Readers in the Nation; for the Strokes of your Letters, being too thin and narrow, hurt the Eye, and he could never read a Line of them without Pain (Straus & Dent, 1907, p.19).

However Franklin goes on to elaborate how he later tricked the Gentleman by showing him one of Caslon’s specimens as if it was the work of Baskerville, and how the Gentleman did not spot the difference.

It is likely that the antipathy towards Baskerville was actually more related to trade jealousy from London, than it was towards his work as such. The notion of the industrial boomtown of Birmingham producing elegant printing was either absurd - or, if true, deeply disturbing. A likely subject for jealousy is also the fact that Baskerville started out as an amateur. He was an outsider with money, he did not need to make a profit to survive, and he could therefore allow himself to spend the extra time necessary to refine his prints, a luxury that was not permitted to his competitors. Thus, critical views, such as those represented above, were therefore more related to his rivals seeing him as a threat than they were to the quality of his fonts.

Didone

Although a lot had changed in arts and politics in the years between Baskerville and the later Didone styles of Bodoni in Italy and the Didots in France, it is interesting that the Didone type-founders did not meet the same level of objection from their contemporaries, not even when moving their designs much
further towards the extremes than Baskerville ever did. The differences in the contemporary reception of the works of Baskerville and Bodoni are quite remarkable.\footnote{Updike sums up Bodoni’s carrier as follows: ‘He was appointed printer to Carlos III of Spain; he received a pension from his son, Carlos IV; he corresponded with Franklin; he was complimented by the Pope; the city of Parma struck a medal in his honour; he obtained a medal for his work at Paris; he received a pension from the Viceroy of Italy; Napoleon gave him another and a larger one, and in short he was a great personage’ (Updike, 1937, vol.2, p.165).}

Although generally popular, the Didots (fig.111) did experience some resentment towards their typefaces. Citizen Sobry explained the reason for his own opposition in 1800, referring to a test carried out by director Etienne Anisson-Duperron of the Imprimerie Royale. In this test, Anisson applied a distance study\footnote{According to Morison (1928) this is the first known example of a legibility investigation.} comparing a page set in a Garamond font with another in a Didot font, and found the Garamond to be legible at several stages after the Didot was no longer to be read (Morison 1928). This led Anisson to refuse to install the new Didone typefaces in the state printing office (Dreyfus 1982). However, as pointed out by Updike (1937), after the Revolution it was Anisson and not the type that perished. Anisson was guillotined and the Didots took over his printing office in the Louvre.

Despite the disapproval from the director of the state printing office, the Didot fonts were quite popular at the time. One reason for this might be related to the success of the Didot point system.\footnote{In 1783 François Ambroise Didot refined the system of comparative body sizes that Fournier first published in 1737. The idea was most likely inspired by the body size system developed by the Académie des Science in 1695.} It appears that the spread of the system among printers enhanced the superiority of the Didot fonts over others that were not based on the same measuring system (Dreyfus, 1982).

The lack of a noticeable opposition towards the Didone typefaces should also be viewed in relation to the movements of the times. As Updike (1937) noted, the upsurge in interest in antiquity, that occurred around 1800, had a major impact on the development of typefaces. People were looking for fonts that symbolized the greatness of the ancient monuments; furthermore, in post-revolutionary France, with its atmosphere of change, the novelty of Didot typefaces became a symbol of the new enlightenment in their mathematical exactness and severity (Nesbitt, 1957). Another reason for the lack of disapproval is supposedly related to the fact that most of the work of both

---

\textbf{Figure 111.} Types of folio Horace by Pierre Didot of 1799 (Updike, vol.1, p.231).
Bodoni and the Didots was rather expensively aiming at serious bibliophiles (Bartram, 2001). It is likely that many of the books functioned as collectors’ items more than as actual, popular reading material, a matter further supported by Bodoni being rather careless when it came to the proofreading of his work (Updike, 1937; Lawson, 2002). He seems to have been more interested in the overall look of the page than in the actual reading matter (fig.112). He was not interested in communicating to the masses, he was a court printer, and his work was aimed at the elite (Updike, 1937). As noted by Haley, Bodoni’s work ‘was probably the most honoured – and least read – printing of his time’ (Haley, 1987, p.14). So by choosing an audience fascinated by the ability to create fine hairlines and clear printing quality, the Didone typefaces ended up being highly regarded. Although later commonly agreed upon that due to the contrast relation of heavy stems and thin hairlines, the style results in low legibility.

**Serif and Sans Serif**

The legibility of the Sans Serif is still a subject for discussion though most readers today would agree that it does not cause them great trouble reading a text set in an ordinary Sans Serif typeface.

Before the Sans Serif became a part of everyday life, traditionalist designers and writers strongly opposed its application in reading material conventionally dominated by the Serif typeface. Stanley Morison fought against this up and coming threat, by stating:

 [… ] the serif is a device which centuries of experience of reading has made into a convention so strong that the efforts of artist, intellectuals and educationalists have never been strong enough to destroy it (Morison, 1959, p.xi).

---


37) For arguments in favour of serif faces over Sans Serif faces see: Updike (1937, vol.2, p.243); Mergenthaler Linotype Company (1935, p.40); Ovink (1938, p.78); Laker (1946, p.13); Tarr (1949, p.31); Deman (1955, p.136 & p.186); Warde (1956, p.55); Burt (1959, p.9); Kindersley (1960, p.463 & p.465); Gray (1960, p.40); Zachrisson (1965, p.36); Robinson et al. (1971, p.359); McLean (1980, p.44); Wallis (1985, p.95); Tracy (1986, p.31); Smeijers (1996, p.32); Goines (1999, p.26); and Kennedy (2003, p.320).
It is remarkable that even at a time where the Sans Serif was rather unpopular amongst traditionalist designers, it performed quite well in test situations. Pyke (1926) found that the only Sans Serif applied in his study ranked second, being 18% less legible than an Old Style font, and with a 30% superiority to the third on the list, a Didone font. In another study carried out by Paterson & Tinker (1932) comparing 10 different fonts including Kabel Light as the only Sans Serif, the performance of Kabel Light was slightly inferior to the Old Style and Didone fonts and superior to the typewriter and Gothic fonts in the test (fig.113). According to the authors, the difference between the Sans Serif
and the Old Style and Didone fonts was not statistically reliable, and they therefore concluded that Kabel Light, the Old Style, and Didone faces were equally legible. However, when studying reader’s preferences of the same fonts, Kabel Light was judged as the second most illegible of all 10 fonts (Tinker, 1964).

In a discussion of the loss of interest in legibility research among contemporary designers, Wim Crouwel refers to the lack of relevance of the research today. His argument is that because of the wide range of media, readers have grown used to the constant transactions. Crouwel goes on to speculate that since reading habits have changed so dramatically, if the former legibility studies were carried out today, the results would probably show no difference between the typefaces tested (Crouwel, 2001). This argument might apply to the comparison of well-known Sans Serif and Serif faces, however, typefaces that are unfamiliar today, such as Gothic and flourishing script faces, will now trouble the reader just as much as unfamiliar typefaces have always done. As Lord Stanhope succinctly put it about 200 years ago:

I can easily conceive that many fantastical flourishes which are given to some letters in old printed books, had their admirers when these books were first published; but it would be difficult for any person now to succeed, if he were to undertake to show their superior beauty, and insist upon their restoration to use (Hansard, 1825, p.477-478).

Familiarity follows from experience, so to be troubled by unfamiliar type one must already be familiar with a certain style of lettering. As demonstrated in the Middle Ages, familiarity with letterforms was an issue for skilled readers alone. Furthermore, a summary of the anecdotal references of the Gothic and Didone typefaces, gives the impression that clarity of letters is a concern mostly of stable societies. In unbalanced environments, the emphasis on aesthetic, social or ideological identity will be viewed as the most significant factors, and legibility will as a result fall into the background. However as the Nazis discovered, a total disregard of readers’ habituation will likely result in the message never reaching the audience. Whereas a quick switch between
Gothic and Latin typefaces is too huge a development to follow for a skilled reader, minor changes to the alphabet, such as the coming of Didone and Sans Serif typefaces, are possible for the reader to cope with as long as they have an open mind and time to adjust.

All the typeface styles mentioned started out as being unfamiliar to the reader, yet in the case of the early scripts, and in the shift between Gothic and Latin alphabets, not only the style of the type but also the letterforms were unfamiliar, a fact that appears to have troubled the Gothic-Latin shift. The work of Baskerville, and the Didone and Sans Serif categories, on the other hand, all contained letterforms that in most cases were commonly known. Although none of these last typefaces avoided criticism from contemporary opinion makers, the public sooner or later accepted the new styles. This suggests that both the level of exposure and the level of common letter features have an influence on familiarity.
3.2

A study of familiarity before and after an exposure session

Prior to a typeface being established in society, familiarisation with the style and letterforms usually happens gradually. Little is known about the changes that take place in the reader’s mind and in the perceptual system as a particular typeface, a typeface style, or new letterforms become more commonly known. Drawing on the findings of the visibility investigation (see Chapter 2.5), the present study investigates how exposure to different typefaces influences the reader.

Sanocki (1988) compared strings of letters from one font with strings of mixed letters from two fonts of different typefaces. Setting out to identify the familiarity effect with a forced choice study and short exposure of the stimuli, Sanocki found that accuracy in performance was higher overall when the letters were from one font alone. He concluded that the results were consistent with the idea that the perceptual system is ‘tuned to the regularities of a particular font in order to process visual information efficiently’ (Sanocki, 1988, p.472).

Another study, carried out by Zineddin and colleagues (2003), looked into improvement on distance threshold for typeface familiarity by testing participants of two different age groups. The
test stimuli consisted of three fonts with different familiarity levels in the letterforms. Improvement in letter and word legibility was compared between one font with an exposure session, and two other fonts without an exposure session. The methodology was as follows: the three fonts were tested on a visual acuity chart; one of the fonts was used for an exposure session of reading text aloud; then a second acuity test was carried out on all three fonts, followed by a second exposure session on the same font used in the first exposure session, followed by a third acuity test on all three fonts.

The authors found a general improvement of performance for the fonts used in the exposure sessions, and established that exposure is part of familiarity. Combining these experimental findings with historical data on how the popularity of certain typefaces have changed, confirms the widely held theory that exposure can transform a typeface’s status among readers. The present study builds on this work by additionally investigating the role of letterforms.

Test material
To identify the features that make a typeface unusual to the reader, Frutiger’s (1998) model of the letterform matrix is useful (fig.114). The idea is that every character has a basic skeleton letter based on a collective memory of all the different character variations a person has ever encountered. To demonstrate this concept, Frutiger superimposed eight of what he defines as the most widely read typefaces. The shared area of these faces represents the matrix of that particular letter. Hypothetically, it is possible to make an area comparison between a letter in a particular font and the average product of multiple popular fonts. This exercise has not been carried out here.

Three test conditions are examined in the present familiarity investigation. The first condition (known-normal) consists of the widely used typefaces forming the base of the Frutiger model. One key member of this group is the typeface Helvetica; as demonstrated in the Helvetica movie (2006), this is probably the most extensively used typeface on the market today. Another member is the typeface Times New Roman. Besides a history of being massively applied in newspapers, books, and other printed materials, it is widely used for office reports due to it being the default font in Word until Microsoft Office 2007.
Other extensively used typefaces are the various Old Style revivals and all faces that we, through exposure, have grown used to over the years. The second condition (new-normal) consists of newly designed typefaces with well-known letter features that fit into the Frutiger model. Because of their recent creation date, these typefaces have yet to be extensively exposed to the reader. The third condition (new-abnormal) consists of typefaces with characters that appear both novel and unusual, with shapes differing from the Frutiger model.

Two studies were carried out with participants in each study conducting tasks with a font from each of the three familiarity conditions. Monotype Originals Times New Roman Regular and Linotype AG Helvetica Regular were chosen for the known-normal condition. Tinker Regular and PykeText Regular with default OpenType settings were chosen for the new-normal condition, and the same two fonts with the Stylistic Set 1 OpenType feature active were chosen for the new-abnormal condition. A high number of existing fonts with little exposure and unusual features also have low visibility. Due to the findings of the visibility investigation, the new-abnormal fonts applied here are known to

Table: Different levels of typeface familiarity.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Group-1</th>
<th>Group-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1) Known-normal</td>
<td>Times New Roman</td>
<td>Helvetica</td>
</tr>
<tr>
<td></td>
<td>- high level of previous</td>
<td></td>
</tr>
<tr>
<td></td>
<td>exposure</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- high level of common</td>
<td></td>
</tr>
<tr>
<td></td>
<td>letter features</td>
<td></td>
</tr>
<tr>
<td>New</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2) New-normal</td>
<td>Tinker</td>
<td>PykeText</td>
</tr>
<tr>
<td></td>
<td>- low level of previous</td>
<td>with default setting</td>
</tr>
<tr>
<td></td>
<td>exposure</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- high level of common</td>
<td>with default setting</td>
</tr>
<tr>
<td></td>
<td>letter features</td>
<td>Referred to as Pyke</td>
</tr>
<tr>
<td>3) New-abnormal</td>
<td>PykeText</td>
<td>Tinker</td>
</tr>
<tr>
<td></td>
<td>- low level of previous</td>
<td>with Stylistic Set 1</td>
</tr>
<tr>
<td></td>
<td>exposure</td>
<td>Referred to as Pyke</td>
</tr>
<tr>
<td></td>
<td>- low level of common</td>
<td></td>
</tr>
<tr>
<td></td>
<td>letter features</td>
<td></td>
</tr>
</tbody>
</table>

38) Examples of popular revivals are: Centaur, ITC Garamond, Adobe Caslon, and Sabon.
39) The second category holds typefaces like Palatino Sans and Meta Serif, both based on standard forms, and both having famous siblings, however at the time of their release in 2005 and 2007, they were unknown to the broader public.
40) The third category holds the now unfamiliar Gothic and flourished script faces, and the various typefaces attempting to reinvent the alphabet, such as ‘basic alphabet’ by Herbert Bayer and ‘New Alphabet’ by Wim Crouwel.
41) The typefaces Tinker and Pyke, both have two forms of OpenType settings. One is the default setting that is based on the familiar letter features of the visibility investigation; the other is the stylistic set 1 feature that incorporates the unfamiliar letter features from the visibility investigation into the typefaces.
1. The older the school pupil becomes, the stronger is the force of those economic and social influences which ultimately will remove him from the school. Up to the age of fourteen the public school hold the pupils well aided by the compulsory attendance laws under the guidance of the grocery clerk.

2. The older the school pupil becomes, the stronger is the force of those economic and social influences which ultimately will remove him from the school. Up to the age of fourteen the public school hold the pupils well aided by the compulsory attendance laws under the guidance of the grocery clerk.

3. The older the school pupil becomes, the stronger is the force of those economic and social influences which ultimately will remove him from the school. Up to the age of fourteen the public school hold the pupils well aided by the compulsory attendance laws under the guidance of the grocery clerk.

1. The older the school pupil becomes, the stronger is the force of those economic and social influences which ultimately will remove him from the school. Up to the age of fourteen the public school hold the pupils well aided by the compulsory attendance laws under the guidance of the grocery clerk.

2. The older the school pupil becomes, the stronger is the force of those economic and social influences which ultimately will remove him from the school. Up to the age of fourteen the public school hold the pupils well aided by the compulsory attendance laws under the guidance of the grocery clerk.

3. The older the school pupil becomes, the stronger is the force of those economic and social influences which ultimately will remove him from the school. Up to the age of fourteen the public school hold the pupils well aided by the compulsory attendance laws under the guidance of the grocery clerk.
have an equal level of visibility as the new-normal fonts applied. A known-abnormal condition was not included due to difficulties in identifying two typefaces that both have the same level of previous exposure as the old-normal fonts and the same level of uncommon letter features as the new-abnormal fonts.

The investigation has two main hypotheses for the outcome. The first (1) is that exposure contributes to familiarity. The findings of the experimental studies mentioned above (Sanocki, 1988; Zineddin et al., 2003) indicate the presence of a font-tuning system. Thus, some improvement in all of the fonts is therefore expected. However, fonts of the known condition will be expected to show a higher reading speed in the pre-test, and fonts of the new conditions, will be expected to improve at a faster rate than the others. The results should generally show a main effect for exposure with more text being read after the exposure session than before the exposure session, and an interaction effect with the new fonts increasing more than the known fonts. The other (2) hypothesis is that common letter features are important, and that the normal condition fonts will perform better than the abnormal condition fonts, both before and after the exposure period. A main effect is expected for the font variable, but no main effect is expected for exposure and an interaction effect is not expected.

Method

The goal was to study the objective level of reading performance of fonts of the different conditions, and the subjective influence these fonts have on the reader. To investigate how fonts of different familiarity conditions behave in relation to exposure, reading speed and readers’ opinion were measured both before and after an exposure session.

Participants

60 participants from the student and staff communities at the Royal College of Art and the Imperial College participated in this study. 30 participants saw the group-1 conditions and a different 30 participants saw the group-2 conditions. Their ages ranged from 20 to 52 with an average age of 28. All participants were native English speakers. They all self-reported having either normal or corrected-to-normal visual acuity, and reported reading
between 1-10 hours per day with an average of 3.3 hours. Participants were compensated with either a gratuity of Microsoft software or £15.

Material
The test material layout was created to present the fonts with a high level of visual similarity. The visibility of the two font conditions of Tinker and the two font conditions of Pyke, were calibrated in the visibility investigation. However, since the tested fonts in general vary on x-height and width, the sizes and leading had to be adjusted internally. Pyke was therefore presented in a 10.3 point size with a leading of 16.5 point, Tinker in a 9 point size with a leading of 16 point, Times in a 10.5 point size with a leading of 17 point, and Helvetica in a 9.3 point size with a leading of 17 point. All test material had a column width of 10.5 cm and was printed on 80 gram white A4 copy paper on a Brother HL-5220 Laser printer.

Procedures
Participants completed tasks for all three typeface conditions; the order in which the typeface conditions were read, and the text applied in the test material were counterbalanced (i.e. all six possible orders were used an equal number of times). Each condition consists of a reading speed pre-test, a pre-test questionnaire, an exposure session, a reading speed post-test, and a post-test questionnaire. Two separate test groups were applied. This was done to avoid the same participants reading both Tinker Default and Tinker Stylistic set 1, or the same participants reading both Pyke Default and Pyke Stylistic set 1. By separating participants in two test groups all new condition fonts are guarantied to be new to the reader.

The questionnaire was designed to examine reader’s opinion not quantified with a speed of reading test. As discussed in chapter 1.4, if focus is performance and visibility, readers’ preferences is non-essential, yet when focus is familiarity the subjective experience plays a central role, and so makes a study into readers’ opinion a useful method.

Participants were presented with six different statements. Four statements were identical in the pre-test and post-test questionnaires and concerned concentration, comprehension, conformability, and future interest in the type. One statement,
which was only presented in the pre-test questionnaire, was concerned with whether participants believed that they had encountered the typeface before, and a last statement – which was only presented in the post-test - was concerned with whether the typeface was easier to read after the exposure session.

The reading speed pre-test consisted of a timed reading test using one of the typeface conditions. The reading material contained a number of short paragraphs, each with a phrase or statement towards the end of the text that stands out from the rest by making the meaning of the paragraph absurd. The task was to identify as many of these phrases as possible within a 2-minute task period. An example of a paragraph follows:

The older the school pupil becomes, the stronger is the force of those economic and social influences which ultimately will remove him from the school. Up to the age of fourteen the public school hold the pupils well aided by the compulsory attendance laws under the guidance of the grocery clerk.\textsuperscript{42}

To reach a correct response, participants were to mark the word 'grocery', after which they moved on to the next paragraph. The paragraphs are structured in such a way that participants must read the whole text to locate the wrong word.

After the reading speed pre-test, participants responded to the pre-test questionnaire. After the pre-tests, the next session was a 20 minutes exposure session where participants read a number of short stories printed in the font under study. The two post-tests followed the exposure session. Participants took a reading speed post-test identical to the reading speed pre-test, but with different short paragraphs. Lastly, following the completion of the reading speed post-test, the participants responded to the post-test questionnaire about their reading comfort. This sequence of tests was repeated for each of the three familiarity conditions with a 5-minute break between each condition.

\textsuperscript{42} Miles A. Tinker, Tinker Speed of Reading Test © 1947, 1955 (renewed 1983) by Miles A. Tinker. Published by the University of Minnesota Press.
Results

Although participants from the Royal College of Art demonstrated a marginally slower reading rate, there was no reliable difference in reading performance between participants from the Royal College of Art or the Imperial College (p>.05). There were also few differences in their questionnaire scores. While it could be assumed that artists and designers would reply to the statements differently than others, this was not the case in the present study. All further analyses will combine the reading performance and questionnaire results from the two schools’ participants.

Participants made very few errors in selecting the out of place word, therefore there was no useful reason to analyze these errors.

Reading speed of group-1

The data shows that the average number of paragraphs read in the known-normal Times during the pre-test was 6.12 (SD = 2.11 paragraphs) and during the post-test 6.80 (SD = 1.70 paragraphs). The average number of paragraphs read in the new-normal Tinker during the pre-test was 6.48 (SD=1.80 paragraphs) and during the post-test 6.42 (SD= 1.80 paragraphs). The average number of paragraphs read in the new-abnormal Pyke during the pre-test was 5.78 (SD = 2.13 paragraphs) and during the post-test 6.52 (SD = 1.85 paragraphs).

![Figure 118. Average number of paragraphs read for group-1.](image-url)
A 2x3 two-way analysis of variance (ANOVA) was used to analyze the reading speed data with two levels for the pre- and post repeated measures and three levels for the font conditions. There was a very large reliable main effect for the pre-test/post-test variable, F(1,29)=15.18, p=.0005. Overall participants took longer to read during the pre-test than during the post-test. There was not a reliable main effect for the font variable, known-normal Times, new-normal Tinker and new-abnormal Pyke, F(2,58)=0.70, p>.05, indicating that the reading speed was similar for the three fonts when averaged over the pre-test and post-test. There was a large, but not statistically reliable interaction effect between fonts and the pre- and post-test repeated measure, F(2,58)=2.51, p=.09. There was a greater increase in reading speed for new-abnormal Pyke than there was for new-normal Tinker.

Further examining the reliable main effect for the pre- and post-test repeated measure, we find that the post-hoc tests show reliable reading speed increase for known-normal Times (t(1)=21.38, p=.03), and new-abnormal Pyke (t(1)=76.33, p=.008), but not for new-normal Tinker (t(1)=0.41, p>.05).

**Reading speed for group-2**

A study of the data from group-2, shows the average number of paragraphs read in the known-normal Helvetica during the pre-test was 5.4 (SD=1.86 paragraphs) and during the post-test 5.9
(SD= 1.74 paragraphs). The average number of paragraphs read in the new-normal Pyke during the pre-test was 5.4 (SD=2.05 paragraphs) and during the post-test 5.6 (SD= 2.13 paragraphs). The average number of paragraphs read in the new- abnormal Tinker during the pre-test was 4.9 (SD=1.69 paragraphs) and during the post-test 5.4 (SD= 1.85 paragraphs).

A 2x3 two-way analysis of variance (ANOVA) was used to analyze the reading speed data of group-2, showing a reliable main effect for the pre-test/post-test variable, \( F(1,29)=13.15, p=.001 \). Overall participants took longer to read during the pre-test than during the post-test. There was not a reliable main effect for the font variable, known-normal Helvetica, new-normal Pyke, and new-abnormal Tinker, \( F(2,58)=1.80, p=.174 \), indicating that the reading speed for the three fonts was similar when averaged over the pre-test and post-test. The interaction effect between the fonts and the pre- and post-test repeated measure was not statistically reliable, \( F(2,58)=0.92, p=.404 \).

Further examining the reliable main effect for the pre- and post-test repeated measure of group-2, we find that the post-hoc tests show reliable reading speed increase for known-normal Helvetica (\( t(29)=2.72, p=.011 \)), and new-abnormal Tinker (\( t(29)=2.89, p=.007 \)), but not for new-normal Pyke (\( t(29)=0.39, p=.385 \)).

**Questionnaires**

The pre- and post-test questionnaires were answered on a 7-point Likert scale. Participants were asked to express their level of agreement with a given statements on a scale between the extremes of ‘I strongly agree’ (+3) and ‘I strongly disagree’ (-3), the average of these responses are presented in the graphs below. As in the reading speed study, a 2x3 two-way analysis of variance (ANOVA) was used to analyze the statements.

[Q1:] For the statement ‘I will enjoy reading this typeface in the future,’ the known-normal and new-normal fonts for both test groups averaged a score of 1.2 both in the pre-test and in the post-test. The new-abnormal fonts increased from the average of -1.5 to -0.3 in group-1, and from -1.6 to -0.5 in group-2. Group-1 showed a reliable main effect for pre-test/post-test \( F(1,29)=6.58, p=.01 \), and a reliable main effect for typeface,
F(2,58)=29.91, p<.0001. There was a reliable interaction effect, F(2,58)=10.13, p<.0001. Group 2 showed a reliable main effect for pre-test/post-test F(1,29)=13.59, p<.001. There was also a reliable main effect for typeface, F(2,58)=42.52, p<.0001. There was a reliable interaction effect, F(2,58)=7.64, p=.001. The results of the two test groups are the same demonstrating big changes for the new-abnormal fonts, and no changes for the other two font conditions.

<table>
<thead>
<tr>
<th></th>
<th>Known-Normal</th>
<th>New-Normal</th>
<th>New-Abnormal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Times</td>
<td>Times</td>
<td>Tinker</td>
<td>Times</td>
</tr>
<tr>
<td>Pyke</td>
<td>Pyke</td>
<td>Tinker</td>
<td>Pyke</td>
</tr>
</tbody>
</table>

**Figure 120.** The average response to the statement: 'I was constantly focusing on the typeface'.

For the statement 'I was constantly focusing on the typeface', the known-normal and new-normal fonts for both test groups averaged a score of -1.5 both in the pre-test and in the post-test. The new-abnormal fonts decreased from the average of 1.3 to 0 in group-1, and from 1.2 to 0.3 in group-2.

**Figure 121.** The average response to the statement: 'I will enjoy reading this typeface in the future'.

<table>
<thead>
<tr>
<th></th>
<th>Known-Normal</th>
<th>New-Normal</th>
<th>New-Abnormal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Times</td>
<td>Times</td>
<td>Tinker</td>
<td>Times</td>
</tr>
<tr>
<td>Pyke</td>
<td>Pyke</td>
<td>Tinker</td>
<td>Pyke</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Known-Normal</th>
<th>New-Normal</th>
<th>New-Abnormal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Times</td>
<td>Times</td>
<td>Tinker</td>
<td>Times</td>
</tr>
<tr>
<td>Pyke</td>
<td>Pyke</td>
<td>Tinker</td>
<td>Pyke</td>
</tr>
</tbody>
</table>
Group-1 showed a reliable main effect for pre-test/post-test $F(1,29)=4.18, p=.05$. There was also a reliable main effect for typeface, $F(2,58)=37.55, p<.0001$. There was a reliable interaction effect, $F(2,58)=8.09, p<.001$. Group 2 showed a reliable main effect for pre-test/post-test $F(1,29)=4.90, p=.03$. There was a reliable main effect for typeface, $F(2,58)=62.02, p<.0001$. There was a reliable interaction effect, $F(2,58)=3.64, p=.03$. The results of the two test groups are the same demonstrating big changes for the new-abnormal fonts, and no changes for the other two font conditions.

Figure 122. The average response to the statement: ‘I still remember most of what I was reading’.

[Q3:] For the statement ‘I still remember most of what I was reading’, the known-normal and new-normal fonts for both test groups averaged a score of 0.2 in the pre-test and increased in the post-test to 0.5. The new-abnormal fonts increased from the average of -0.8 to 0 in group-1, and from -0.3 to 0.5 in group-2. Group 1 showed a reliable main effect for pre-test/post-test $F(1,29)=5.58, p=.025$. There was also a reliable main effect for typeface, $F(2,58)=9.42, p<.001$. There was not a reliable interaction effect, $F(2,58)=3.02, p=.056$. Group 2 showed a reliable main effect for pre-test/post-test $F(1,29)=12.98, p=.001$. There was not a reliable main effect for typeface, $F(2,58)=1.67, p=.20$. There was not a reliable interaction effect, $F(2,58)=1.52, p=.23$. The effect size and differences are small in both test groups, the strongest effect is found in the new-abnormal conditions.
[Q4:] For the statement 'This was a comfortable reading experience’, the known-normal and new-normal fonts for both test groups averaged a score of 0.9 in the pre-test and in increased in the post-test 1.3. The new-abnormal fonts decreased from the average of -1 to 0 in group-1, and from -1.4 to -0.1 in group-2. Group 1 showed a reliable main effect for pre-test/post-test $F(1,29)=10.3$, $p=.003$. There was also a reliable main effect for typeface, $F(2,58)=18.78$, $p<.0001$. There was a reliable interaction effect, $F(2,58)=3.35$, $p=.04$. Group 2 showed a reliable main effect for pre-test/post-test $F(1,29)=27.43$, $p<.0001$. There was also a reliable main effect for typeface, $F(2,58)=21.59$, $p<.0001$. There was a reliable interaction effect, $F(2,58)=3.67$, $p=.03$. In both test groups there is a small increase from pre to the post-test for the new-normal conditions as well as a large increase for the new-abnormal condition.

Figure 123. The average response to the statement: 'This was a comfortable reading experience'.
Figure 124. The average response to the statement: 'I have encountered this typeface before'.

Figure 125. The average response to the statement: 'I find the typeface easier to read now, than I did at the beginning of the test'.

[Q5.1:] For the statement 'I have encountered this typeface before', (only asked in the pre-test), the known-normal and new-normal fonts for both test groups averaged a score of 1.4. The new-abnormal fonts average a score of -2.6 in group-1, and from -2.3 in group-2. A one-way ANOVA for group-1 showed a reliable difference for this question, $F(58) = 76.9$, $p < .0001$. Post-hoc tests showed that all three conditions were reliably different from each other. A one-way ANOVA for group-2 showed a reliable difference for this question, $F(58) = 149.5$, $p < .0001$. Post-hoc tests showed that new-abnormal Tinker was reliably different from both known-normal Helvetica and new-normal Pyke.

The results of the two test groups are the same in known-normal and new-abnormal fonts, and different in the new-normal condition fonts.
For the statement 'I find the typeface easier to read now, than I did at the beginning of the test' (only asked in the post-test), the known-normal and new-normal fonts for both test groups averaged a score of 0.8. The new-abnormal fonts average a score of 1.9 in group-1, and from 1.1 in group-2. A one-way ANOVA for group-1 showed a reliable difference for this question, $F(58) = 9.73, p=.0002$. Post-hoc tests showed that new-abnormal Pyke was reliably different from both known-normal Times and new-normal Tinker. A one-way ANOVA for group-2 did not show a reliable difference for this question, $F(58) = 2.27, p=.11$.

**Discussion**

Reading speed increased from the pre-test to the post-test for four of the six fonts. Whereas the improvement in the new-normal conditions either did not happen or was small, the known-normal and new-abnormal conditions in both test groups increased in reading speed after the exposure session. There were, in addition, a noticeably high number of reliable differences in the questionnaire data with a great similarity between the two groups, showing that readers’ subjective impression of new-abnormal condition fonts improves through the short practice, and that readers do not detect a difference in the reading of known-normal and new-normal condition fonts.

**The exposure hypothesis**

The hypothesis that familiarity improves by exposure had some support in the reading speed study. The data demonstrated a main effect for exposure in both test groups, where participants read at a reliably faster rate after the exposure session. But the expected interaction effect was not reliable in either of the groups. The interaction showed that the two largest improvements from exposure came for the known-normal and new-abnormal conditions, when it was expected that the largest improvements would come of the two new conditions. The improvement in the known-normal fonts indicates the possibility that the exposure effect in reading performance can be driven by short-term font-tuning rather than long-term repeated viewings over a lifetime. It appears that the fonts tested in the new-normal condition
are less susceptible to font-tuning than the known-normal fonts tested. If this is the case we can conclude that the tested new-abnormal fonts – originating in the same typefaces as the tested new-normal fonts – are also not influenced by font-tuning and therefore must be influenced by practice alone.

The expected difference between the pre-test normal conditions was not confirmed, showing no reliable interaction effect in either reading speed or readers’ opinion. If previous exposure with a particular typeface is important, then the known-normal fonts should be superior to the new-normal fonts in the pre-test. The findings suggest that the level of previous exposure with a particular typeface is not a factor for the reader, and that new typefaces have no negative influence on the reading process. The investigation also shows that exposure can have a significant impact on readers opinion of new-abnormal condition fonts. Being highly negative in the pre-test questionnaire, readers found in the post-test that comprehension went up, that they focused less on the shapes of the letters, that their level of reading comfort went up, and that they found themselves being more positive towards reading the typeface in the future. This notable change of opinion between the pre- and post-tests were not similarly present in the normal condition fonts, which in most accounts delivered rather identical answers both before and after the exposure session.

The letter feature hypothesis

The hypothesis that common letter features are important to reading was confirmed in the pre-test, demonstrating a better performance for the two normal conditions fonts compared to the abnormal condition fonts. This is further supported by the questionnaire data strongly supporting the letter feature hypothesis, on all accounts showing the new-abnormal condition to be less preferred than the normal conditions.

However, after the exposure session the abnormal font was read as fast as the normal fonts in both test groups, a remarkable finding demonstrating that when visibility is accounted for, readers will quickly adjust themselves to letters highly different from the Frutiger model, and so reach an equal reading speed as fonts of common letter features and high previous exposure.
The findings of the reading speed study are unexpected. We see a main effect for exposure confirming the belief that exposure has a strong effect on reading. Yet, we see no main effect for the different fonts tested. This revelation that unusual letterforms do not slow down reading after a brief exposure period, surprisingly tells us that the level of common letterforms in typefaces is not important to reading performance. The results go against the argument put forward by Stanley Morison and many other typographers who emphasises that, for a successful reading experience, letterforms should stay constant. So, if typefaces of uncommon letter features are read equally as fast as typefaces of common letter features, why is it that letter skeletons have changed so little over the years? The answer appears to lie in readers' subjective experience. The findings of readers' opinion follow the view of Morison and not the findings of the reading speed study. While the collective reading performance of each font did not differ from the other fonts tested, readers were noticeably more critical towards the fonts of uncommon letter features compared to the fonts of common letter features. Historical references further shows that a shift between Gothic and Latin alphabets causes problems for the reader; the present investigation indicates that these difficulties do not originate in reading performance. If in fact readers' opinion is the sole influence on typeface familiarity, one could argue for a less conservative approach to the design of new typefaces.
Section 4

CLOSURE
4.1.

Summary, perspectives & conclusion

As Lucien Alphonse Legros argued (1922), legibility appears to be influenced both by objective and subjective factors. The two studies of the visibility investigation are based on objective performance based tests. The familiarity investigation, on the other hand, contains both subjective and objective elements. The subjective questionnaires were designed to study readers’ experience of the reading process, where the more objective reading speed study was designed to measure performance. By focusing both on objective and subjective elements, matters of typeface legibility have been studied from different angles, and so provided interesting new findings.

The data of the visibility investigation suggest ways of making small improvements to familiar letter features and further identify a number of unfamiliar letter features with a similar level of visibility. Both studies of the visibility investigation were carried out on screen. To minimize the influence of resolution and backlit display, the test material was presented in a rather
large size and on a background of a toned down colour. Chapter 2.1 shows that although many issues of visibility are related directly to the specific application, others can be transformed across platforms, a notion suggesting that findings can be valid in relation to media other than the computer screen. The visibility investigation showed that 1) most letters of narrow forms benefit from being slightly broadened, 2) the middle part of the letters ‘a’ and ‘s’ will profit from a round movement instead of diagonal, 3) the crossbar of the ‘e’ should be at the visual centre, 4) a closed aperture of the ‘c’ is just as legible as an open aperture in the parafoveal view, 5) a serif on top of the stem will improve the distance viewing of the letter ‘i’, 6) that unfamiliar extending versions of ‘a’ and ‘s’ are as visible as the familiar versions, and (7) so are an x-height uppercase ‘n’ and an x-height uppercase ‘t’ in the parafoveal view. This investigation was a necessary step along the way to control the visibility of the new conditions in the later familiarity investigation.

The range of known components that influence typeface visibility is rather extended, however, the same cannot be said for typeface familiarity. Although it is commonly agreed among designers that familiarity is of most importance, we have very little knowledge of the different issues influencing this phenomenon. Historical references demonstrate that readers find it difficult to adjust to significant changes in letterforms, and that, in the long run, they are more adaptable to minor changes of style. The current research adds to this body of knowledge by investigating how these adjustments function in the initial stages.

The familiarity investigation found that, 8) immediate exposure to a font influences reading performance, 9) the level of uncommon letter features is nonessential in relation to reading performance, yet essential in relation to reading experience, 10) fonts of uncommon letter features improve in both reading speed and reader’s opinion after a short practice, 11) previous exposure to a font of common letter features has no influence on either performance or opinion.

While these findings clearly indicate the presence of a familiarity effect, the effect did not match perfectly any of the hypotheses, leaving new questions for study.
Future familiarity studies
The familiarity investigation, propose the presence of a font-tuning system. But why does font-tuning not work on all fonts? While great effort was taken to control the visibility between the low exposure default and OpenType stylistic faces in the familiarity investigation, the high exposure faces Helvetica and Times were not included in the visibility investigation. One possibility for the improvement of Helvetica and Times is that the visibility is lower in these faces than in the low exposure faces, and that the reader had adjusted to the visibility during the exposure period and not familiarity. So to exclude visibility as an influence, all categories would benefit from being subjected to a previous visibility screening.

The familiarity investigation further showed that fonts of uncommon letter features after a short exposure reaches an equal level of reading speed as fonts of letter features highly known to the reader. What is the longitudinal influence of this, will the improvement of the uncommon letter features stay with the reader if the font is tested again a week later? Or will the reader have to practice all over again?

In videogame theory the term ‘immersion’ describes the state of mind in which the player loses the sense of a physical self and moves to a level of intense focus and effortless action (Varney, 2006), as it is often pointed out by typographers this frame of mind, called the ‘Disappearing Types’ by Gerard Unger (Unger, 2007, p.45), is also achievable in reading. In the familiarity investigation, the closest participants came to immersive reading was in the 20 minutes exposure sessions. In contrast to the pre-test, the post-test followed directly after an exposure session, and would therefore likely be influenced by immersion. What happens before immersive reading takes over? It might be that various styles of typefaces and familiarity levels influence immersive reading in different ways. It will be worthy of note to compare the present findings with a study where the post-test was separated from the immersive reading condition by a break after the exposure sessions.

The data of the familiarity investigation further motivates a study into an alternative hypothesis suggesting that familiarity is biased towards older, more well-known typeface styles compared to styles of a newer date. The answers to the questionnaire defined
new-normal Tinker to be less familiar than the known-normal font, while new-normal Pyke was defined to be equally familiar as the known-normal font. The reading speed data of both test-groups further showed a lower increase in the Tinker fonts than in the other fonts. This might suggest that newer typeface styles takes longer to improve in reading speed than older ones. It would be interesting to apply the methodology of the familiarity investigation to a number of fonts of newer typeface styles, and compare these with a number of fonts of older typeface styles.

Future visibility studies
The visibility of a typeface is obviously related to more than differentiated roman lowercase characters; the clarity of uppercase, italic and digits are also important matters. Not to mention the range of other variables that would all benefit from a thorough investigation in order to supplement existing knowledge.

Starting from the assumptions of the Parallel Letter Recognition model, the focus of the present visibility investigation was on the bottom-up effect of the individual letter, and not on the top-down effect of the word and the lexical library. To get an accurate understanding of the influence letter shape has on reading, a thorough study into letter groups and typeface rhythm would be beneficial. As expressed by Matthew Carter:

> The central paradox of type design is that in an immediate sense we design letterforms, but letterforms are not our product. We are really word-shape designers; it is only in combination that letters become type (M. Carter, 1995, p.186).

Beatrice Warde strongly advocated the idea that type should be ‘studied in groups of words first of all, then by single words, and last of all by the structure of each letter’, her point being that a lowercase ‘h’ has no reason for existence unless it forms ‘a perfect and matter-of-fact connexion’ with the preceding and following characters (Warde, 1927, p.4).

The spacing of the letters can influence character visibility tremendously. If a text is too closely set, the characters will crowd and their individuality will dissolve, if, on the other hand, the typeface is too widely spaced, the word picture will vanish.
and the reader will be forced to read the letters serially instead of parallel. With the OpenType technology, it is possible to automatically apply contextual alternates designed to avoid bridging in difficult character combinations. An interesting study would be to investigate what kind of influence these extra character shapes have on the reading process.

The visibility investigation found no difference in the low contrast letter ‘c’ having open or closed apertures, but what happens when the ‘c’ is presented in combination with other letters in a word? If the character is placed to the left of a stem, will the stem visually close up the counter by eating the white space, or will the small aperture actually prevent a bridging between the ‘c’ and the stem?

Ole Lund (1999) found no general proof for favouring either Serif or Sans Serif styles, however Harris (1973) demonstrated that serifs seem to improve visibility on certain letters and reduce it on others, and Morris & colleagues (2002) found that a Slab Serif version of Lucida, in small point sizes slow down reading during Rapid Serial Visual Presentations compared to a modification of Lucida without serifs. It appears that future research could benefit from looking into the different kinds of serifs, and study the influence they have on various letterforms.

As discussed in Chapter 2.1, width, weight, and contrast are highly influential factors on type at distance. Defining the ideal values for these parameters for a given signage typeface, both in positive and negative colour, is another central subject for investigation.

While discussing the proportion of the lowercase characters, Fournier suggested in his Manuel typographique from 1764, a relationship where the x-height is of 3 units (about 43%) and the ascender and descender of 2 units each (H. Carter, 1930). Sumner Stone (1989) also advocated a relative low x-height, stating that tall ascenders tend to emphasize distinctness of letter and word shapes, and in that way enhance legibility. Referring to the phenomenon that the upper half of lowercase letters are read more easily than the lower half, Harry Carter (1937) suggested a solution where descending elements were kept short and ascending elements long. This view is supported by Allen Hutt, an authority in newspaper design, who suggested that the letter proportion should be so that the height to width relation-
ship is 'oblong, not square', and that the x-height area should avoid encroaching on the ascenders, arguing that the ascending elements have an important optical function (Hutt 1967). Walter Tracy also emphasised that if the x-height becomes too big the extenders vanish and the individuality of characters reduces (Tracy, 1986).

Nonetheless, typefaces created either for setting small text sizes, or for generally saving space on the printed page, often have a large x-height. One of these typefaces is Demo with an x-height of about 55%. According to its designer, Gerard Unger (1979), a large x-height often requires more leading in setting the text and, as a consequence, actually nullifies the saved space; yet the argument in favour of the high x-height would suggest that most of the essential letter features are to be found in this area anyway, and therefore that an enlargement of these features enhances the visibility. The view is supported by a reading test comparing newspaper pages with bad inking, carried out in connection with the development of the typeface Edison by Hermann Zapf (1987); this resulted in a large x-height and a reduced size of the uppercase letters compared to the ascending lowercase characters. These opposing opinions on letter proportion, motivates a future investigation of the influence of different x-heights in different media.

The visibility investigation was carried out in a large size on screen – quite different from the sizes normally applied in text for continuous reading. Although the anecdotal references presented in Chapter 2.1 show that some visibility factors are identical across platforms, other factors are not. It would be relevant to empirically investigate whether the findings of the visibility investigation would be similar if the test material were printed on paper and/or with different font sizes.

Research comparing printed material and text on screen was quite a popular subject during the 80s (for summaries see: Mills & Weldon, 1987; Dillon et al., 1988), however a number of these studies compared negative presentation of dot matrix screen fonts with positive presentation on paper – an approach leaving

43) According to Tracy this issue is more evident with regards to the ascending characters ‘b’, ‘d’, ‘h’, ‘k’, and ‘I’, than in the descending characters ‘g’, ‘j’, ‘p’, ‘q’ and ‘y’. Tracy advocates an internal relation where the x-height should be about six to ten of the ascending character ‘h’.
too many variables unaccounted for, and not of much use today due to technical developments. Since then, several examples of empirical research on reading typefaces from screen have been put forward, such as the various investigations into the effect of Microsoft’s ClearType technology (Chaparro et al., 2006; Dillon et al., 2006; Sheedy et al., 2008). However as pointed out by Dyson (2004), the body of knowledge on the influence the screen has on typeface design is rather limited, a fact that is partly due to the rapid development of technology, making it difficult to extend to findings just a few years old. Several studies have shown that the higher the quality of on-screen material, the better the performance will be in general (Wright, 1999; Sheedy et al., 2005; Dillon et al., 2006). Screen resolution on most laptops and desktops is now about 100 pixel per inch – much more than it was 20 years ago – an improvement that has had a significant influence on the visual presentation of on-screen material, yet the resolution is still not at a stage where most readers feel comfortable reading longer pieces of text on the computer screen (Larson, 2007). A relevant investigation is to define the specific features in typefaces that enhance their visibility on the constantly improving LCD display devices.

The topics mentioned here represent a small collection of unanswered questions that emerged doing the present research; they should be viewed more as the tip of the iceberg, than as an attempt to create a complete list of unanswered legibility related questions. The fact that we, with all our digital technology today, can control variables that would have been difficult at the time of Miles A. Tinker and his contemporaries makes it possible for us to investigate issues that previous researchers would have viewed as too complicated for study. Our lack of knowledge is therefore still rather significant, and as a result the questions answered by this project have generated a new list of questions that would benefit from further investigations.

The practice part of the project is the creation of the three new typeface families. Preliminary versions were applied in the test material of the experimental studies; the results of the studies were then applied in the final typeface families. In that way the
practice work not only plays a role in controlling variables of the test material, but also demonstrates, in a visual language, how to implement legibility improvements in the design of new typefaces. Combined with the anecdotal references of designers, this approach is directly applicable in commercial type designs.

Through a method of practice-based academic research, I have suggested answers to some legibility related questions. The project combines knowledge from the fields of vision psychology, type history and type design. My own background in design has obviously influenced the angle of the work. However, while searching for new knowledge, the design community tends to seek larger overall and more definite answers. In the same search, the scientific community aims at answering smaller questions one by one. While attempting to solve a tiny fraction of the wide range of legibility issues, my investigation has clearly been motivated by the approach of scientists. Many design researchers have tried to identify a universal set of legibility rules with less than great success. No single guideline can be created to solve all the issues. To expand on existing knowledge, one often has to work on the borderline of what is already known. To be able to focus on legibility, which is interconnected with so many indeterminate topics, some assumptions have to be made to move forward. This is an investigation based on a structure of open-ended enquiries, leaving further questions for future research to be built upon.
Data of visibility study

APPENDIX A
### Study 1: TinkerTest Exposure

|   | a | a | a | a | c | c | c | e | e | e | e | e | e | n | n | n | s | s | u | f | f | f | i | i | i | j | j | l | l | t | t | t |
| a1| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| a2| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| a3| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| a4| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| c1| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| c2| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| c3| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| c4| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| e1| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| e2| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| e3| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| e4| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| e5| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| n1| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| n2| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| n3| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| n4| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| s1| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| s2| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| s3| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| u1| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| u2| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| f1| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| f2| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| f3| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| i1| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| i2| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| i3| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| j1| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| j2| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| l1| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| l2| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| t1| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| t2| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| t3| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

- ○: One identification
- ●: Full identification
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

**Study 1: OvinkTest Exposure**

| Study | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |
|       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

**One identification**

**Full identification**
## Study 1: PykeTest Exposure

<table>
<thead>
<tr>
<th>a1</th>
<th>a2</th>
<th>a3</th>
<th>a4</th>
<th>a5</th>
<th>a6</th>
<th>a7</th>
<th>a8</th>
<th>a9</th>
<th>a10</th>
<th>a11</th>
<th>a12</th>
<th>a13</th>
<th>a14</th>
<th>a15</th>
<th>a16</th>
<th>a17</th>
<th>a18</th>
<th>a19</th>
<th>a20</th>
</tr>
</thead>
<tbody>
<tr>
<td>b1</td>
<td>b2</td>
<td>b3</td>
<td>b4</td>
<td>b5</td>
<td>b6</td>
<td>b7</td>
<td>b8</td>
<td>b9</td>
<td>b10</td>
<td>b11</td>
<td>b12</td>
<td>b13</td>
<td>b14</td>
<td>b15</td>
<td>b16</td>
<td>b17</td>
<td>b18</td>
<td>b19</td>
<td>b20</td>
</tr>
<tr>
<td>c1</td>
<td>c2</td>
<td>c3</td>
<td>c4</td>
<td>c5</td>
<td>c6</td>
<td>c7</td>
<td>c8</td>
<td>c9</td>
<td>c10</td>
<td>c11</td>
<td>c12</td>
<td>c13</td>
<td>c14</td>
<td>c15</td>
<td>c16</td>
<td>c17</td>
<td>c18</td>
<td>c19</td>
<td>c20</td>
</tr>
<tr>
<td>d1</td>
<td>d2</td>
<td>d3</td>
<td>d4</td>
<td>d5</td>
<td>d6</td>
<td>d7</td>
<td>d8</td>
<td>d9</td>
<td>d10</td>
<td>d11</td>
<td>d12</td>
<td>d13</td>
<td>d14</td>
<td>d15</td>
<td>d16</td>
<td>d17</td>
<td>d18</td>
<td>d19</td>
<td>d20</td>
</tr>
<tr>
<td>e1</td>
<td>e2</td>
<td>e3</td>
<td>e4</td>
<td>e5</td>
<td>e6</td>
<td>e7</td>
<td>e8</td>
<td>e9</td>
<td>e10</td>
<td>e11</td>
<td>e12</td>
<td>e13</td>
<td>e14</td>
<td>e15</td>
<td>e16</td>
<td>e17</td>
<td>e18</td>
<td>e19</td>
<td>e20</td>
</tr>
<tr>
<td>f1</td>
<td>f2</td>
<td>f3</td>
<td>f4</td>
<td>f5</td>
<td>f6</td>
<td>f7</td>
<td>f8</td>
<td>f9</td>
<td>f10</td>
<td>f11</td>
<td>f12</td>
<td>f13</td>
<td>f14</td>
<td>f15</td>
<td>f16</td>
<td>f17</td>
<td>f18</td>
<td>f19</td>
<td>f20</td>
</tr>
<tr>
<td>g1</td>
<td>g2</td>
<td>g3</td>
<td>g4</td>
<td>g5</td>
<td>g6</td>
<td>g7</td>
<td>g8</td>
<td>g9</td>
<td>g10</td>
<td>g11</td>
<td>g12</td>
<td>g13</td>
<td>g14</td>
<td>g15</td>
<td>g16</td>
<td>g17</td>
<td>g18</td>
<td>g19</td>
<td>g20</td>
</tr>
<tr>
<td>h1</td>
<td>h2</td>
<td>h3</td>
<td>h4</td>
<td>h5</td>
<td>h6</td>
<td>h7</td>
<td>h8</td>
<td>h9</td>
<td>h10</td>
<td>h11</td>
<td>h12</td>
<td>h13</td>
<td>h14</td>
<td>h15</td>
<td>h16</td>
<td>h17</td>
<td>h18</td>
<td>h19</td>
<td>h20</td>
</tr>
<tr>
<td>i1</td>
<td>i2</td>
<td>i3</td>
<td>i4</td>
<td>i5</td>
<td>i6</td>
<td>i7</td>
<td>i8</td>
<td>i9</td>
<td>i10</td>
<td>i11</td>
<td>i12</td>
<td>i13</td>
<td>i14</td>
<td>i15</td>
<td>i16</td>
<td>i17</td>
<td>i18</td>
<td>i19</td>
<td>i20</td>
</tr>
<tr>
<td>j1</td>
<td>j2</td>
<td>j3</td>
<td>j4</td>
<td>j5</td>
<td>j6</td>
<td>j7</td>
<td>j8</td>
<td>j9</td>
<td>j10</td>
<td>j11</td>
<td>j12</td>
<td>j13</td>
<td>j14</td>
<td>j15</td>
<td>j16</td>
<td>j17</td>
<td>j18</td>
<td>j19</td>
<td>j20</td>
</tr>
<tr>
<td>k1</td>
<td>k2</td>
<td>k3</td>
<td>k4</td>
<td>k5</td>
<td>k6</td>
<td>k7</td>
<td>k8</td>
<td>k9</td>
<td>k10</td>
<td>k11</td>
<td>k12</td>
<td>k13</td>
<td>k14</td>
<td>k15</td>
<td>k16</td>
<td>k17</td>
<td>k18</td>
<td>k19</td>
<td>k20</td>
</tr>
<tr>
<td>l1</td>
<td>l2</td>
<td>l3</td>
<td>l4</td>
<td>l5</td>
<td>l6</td>
<td>l7</td>
<td>l8</td>
<td>l9</td>
<td>l10</td>
<td>l11</td>
<td>l12</td>
<td>l13</td>
<td>l14</td>
<td>l15</td>
<td>l16</td>
<td>l17</td>
<td>l18</td>
<td>l19</td>
<td>l20</td>
</tr>
<tr>
<td>m1</td>
<td>m2</td>
<td>m3</td>
<td>m4</td>
<td>m5</td>
<td>m6</td>
<td>m7</td>
<td>m8</td>
<td>m9</td>
<td>m10</td>
<td>m11</td>
<td>m12</td>
<td>m13</td>
<td>m14</td>
<td>m15</td>
<td>m16</td>
<td>m17</td>
<td>m18</td>
<td>m19</td>
<td>m20</td>
</tr>
<tr>
<td>n1</td>
<td>n2</td>
<td>n3</td>
<td>n4</td>
<td>n5</td>
<td>n6</td>
<td>n7</td>
<td>n8</td>
<td>n9</td>
<td>n10</td>
<td>n11</td>
<td>n12</td>
<td>n13</td>
<td>n14</td>
<td>n15</td>
<td>n16</td>
<td>n17</td>
<td>n18</td>
<td>n19</td>
<td>n20</td>
</tr>
<tr>
<td>o1</td>
<td>o2</td>
<td>o3</td>
<td>o4</td>
<td>o5</td>
<td>o6</td>
<td>o7</td>
<td>o8</td>
<td>o9</td>
<td>o10</td>
<td>o11</td>
<td>o12</td>
<td>o13</td>
<td>o14</td>
<td>o15</td>
<td>o16</td>
<td>o17</td>
<td>o18</td>
<td>o19</td>
<td>o20</td>
</tr>
<tr>
<td>q1</td>
<td>q2</td>
<td>q3</td>
<td>q4</td>
<td>q5</td>
<td>q6</td>
<td>q7</td>
<td>q8</td>
<td>q9</td>
<td>q10</td>
<td>q11</td>
<td>q12</td>
<td>q13</td>
<td>q14</td>
<td>q15</td>
<td>q16</td>
<td>q17</td>
<td>q18</td>
<td>q19</td>
<td>q20</td>
</tr>
<tr>
<td>r1</td>
<td>r2</td>
<td>r3</td>
<td>r4</td>
<td>r5</td>
<td>r6</td>
<td>r7</td>
<td>r8</td>
<td>r9</td>
<td>r10</td>
<td>r11</td>
<td>r12</td>
<td>r13</td>
<td>r14</td>
<td>r15</td>
<td>r16</td>
<td>r17</td>
<td>r18</td>
<td>r19</td>
<td>r20</td>
</tr>
<tr>
<td>s1</td>
<td>s2</td>
<td>s3</td>
<td>s4</td>
<td>s5</td>
<td>s6</td>
<td>s7</td>
<td>s8</td>
<td>s9</td>
<td>s10</td>
<td>s11</td>
<td>s12</td>
<td>s13</td>
<td>s14</td>
<td>s15</td>
<td>s16</td>
<td>s17</td>
<td>s18</td>
<td>s19</td>
<td>s20</td>
</tr>
<tr>
<td>t1</td>
<td>t2</td>
<td>t3</td>
<td>t4</td>
<td>t5</td>
<td>t6</td>
<td>t7</td>
<td>t8</td>
<td>t9</td>
<td>t10</td>
<td>t11</td>
<td>t12</td>
<td>t13</td>
<td>t14</td>
<td>t15</td>
<td>t16</td>
<td>t17</td>
<td>t18</td>
<td>t19</td>
<td>t20</td>
</tr>
<tr>
<td>u1</td>
<td>u2</td>
<td>u3</td>
<td>u4</td>
<td>u5</td>
<td>u6</td>
<td>u7</td>
<td>u8</td>
<td>u9</td>
<td>u10</td>
<td>u11</td>
<td>u12</td>
<td>u13</td>
<td>u14</td>
<td>u15</td>
<td>u16</td>
<td>u17</td>
<td>u18</td>
<td>u19</td>
<td>u20</td>
</tr>
<tr>
<td>v1</td>
<td>v2</td>
<td>v3</td>
<td>v4</td>
<td>v5</td>
<td>v6</td>
<td>v7</td>
<td>v8</td>
<td>v9</td>
<td>v10</td>
<td>v11</td>
<td>v12</td>
<td>v13</td>
<td>v14</td>
<td>v15</td>
<td>v16</td>
<td>v17</td>
<td>v18</td>
<td>v19</td>
<td>v20</td>
</tr>
<tr>
<td>w1</td>
<td>w2</td>
<td>w3</td>
<td>w4</td>
<td>w5</td>
<td>w6</td>
<td>w7</td>
<td>w8</td>
<td>w9</td>
<td>w10</td>
<td>w11</td>
<td>w12</td>
<td>w13</td>
<td>w14</td>
<td>w15</td>
<td>w16</td>
<td>w17</td>
<td>w18</td>
<td>w19</td>
<td>w20</td>
</tr>
<tr>
<td>x1</td>
<td>x2</td>
<td>x3</td>
<td>x4</td>
<td>x5</td>
<td>x6</td>
<td>x7</td>
<td>x8</td>
<td>x9</td>
<td>x10</td>
<td>x11</td>
<td>x12</td>
<td>x13</td>
<td>x14</td>
<td>x15</td>
<td>x16</td>
<td>x17</td>
<td>x18</td>
<td>x19</td>
<td>x20</td>
</tr>
<tr>
<td>y1</td>
<td>y2</td>
<td>y3</td>
<td>y4</td>
<td>y5</td>
<td>y6</td>
<td>y7</td>
<td>y8</td>
<td>y9</td>
<td>y10</td>
<td>y11</td>
<td>y12</td>
<td>y13</td>
<td>y14</td>
<td>y15</td>
<td>y16</td>
<td>y17</td>
<td>y18</td>
<td>y19</td>
<td>y20</td>
</tr>
<tr>
<td>z1</td>
<td>z2</td>
<td>z3</td>
<td>z4</td>
<td>z5</td>
<td>z6</td>
<td>z7</td>
<td>z8</td>
<td>z9</td>
<td>z10</td>
<td>z11</td>
<td>z12</td>
<td>z13</td>
<td>z14</td>
<td>z15</td>
<td>z16</td>
<td>z17</td>
<td>z18</td>
<td>z19</td>
<td>z20</td>
</tr>
</tbody>
</table>
## Study 2: TinkerTest Distance

|   | a | a | a | d | c | c | c | e | e | e | e | E | n | n | N | s | s | u | u | f | f | f | f | i | i | i | i | j | j | l | l | t | t | t | T |
| a |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| b |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| c |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| d |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| e |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| f |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| g |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| h |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| i |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| j |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| k |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| l |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| m |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| n |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| o |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| p |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| q |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| r |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| s |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| t |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| u |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| v |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| w |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| x |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| y |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| z |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Study 2: OvinkTest Distance

|   | a | a | a | a | c | c | c | e | e | e | e | E | n | n | N | s | s | s | u | u | f | f | f | i | i | i | j | j | j | l | l | l | t | t | t | t | t | t | t | t |
| a1|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| a2|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| a3|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| a4|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| b1|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| b2|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| b3|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| c1|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| c2|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| c3|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| c4|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| e1|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| e2|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| e3|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| e4|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| e5|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| e6|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| n1|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| n2|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| n3|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| n4|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| s1|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| s2|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| s3|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| s4|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| s5|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| s6|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| s7|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| s8|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| s9|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| s10|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

○ one identification  ● full identification
Study 2: PykeTest Distance

|   | a | a | a | a | a | a | a | a | a | a | a | a | a |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | o | o | o | o | o | o | o | o | o | o | o | o | o |
| b | o | o | o | o | o | o | o | o | o | o | o | o | o |
| c | o | o | o | o | o | o | o | o | o | o | o | o | o |
| d | o | o | o | o | o | o | o | o | o | o | o | o | o |
| e | o | o | o | o | o | o | o | o | o | o | o | o | o |
| f | o | o | o | o | o | o | o | o | o | o | o | o | o |
| g | o | o | o | o | o | o | o | o | o | o | o | o | o |
| h | o | o | o | o | o | o | o | o | o | o | o | o | o |
| i | o | o | o | o | o | o | o | o | o | o | o | o | o |
| j | o | o | o | o | o | o | o | o | o | o | o | o | o |
| k | o | o | o | o | o | o | o | o | o | o | o | o | o |
| l | o | o | o | o | o | o | o | o | o | o | o | o | o |
| m | o | o | o | o | o | o | o | o | o | o | o | o | o |
| n | o | o | o | o | o | o | o | o | o | o | o | o | o |
| o | o | o | o | o | o | o | o | o | o | o | o | o | o |
| p | o | o | o | o | o | o | o | o | o | o | o | o | o |
| q | o | o | o | o | o | o | o | o | o | o | o | o | o |
| r | o | o | o | o | o | o | o | o | o | o | o | o | o |
| s | o | o | o | o | o | o | o | o | o | o | o | o | o |
| t | o | o | o | o | o | o | o | o | o | o | o | o | o |
| u | o | o | o | o | o | o | o | o | o | o | o | o | o |
| v | o | o | o | o | o | o | o | o | o | o | o | o | o |
| w | o | o | o | o | o | o | o | o | o | o | o | o | o |
| x | o | o | o | o | o | o | o | o | o | o | o | o | o |
| y | o | o | o | o | o | o | o | o | o | o | o | o | o |
| z | o | o | o | o | o | o | o | o | o | o | o | o | o |

○ one identification  ● full identification
Tinker

**APPENDIX B**
The typeface **Tinker**, named after the legibility researcher Miles A. Tinker, is a hybrid created for a high differentiation of characters. An early version of the typeface was subjected to experimental legibility investigations of distance and time threshold methods. With the typeface Tinker, participants were exposed to different variations of the most frequently misread lowercase letters. The findings, which were implemented in the final designs, demonstrated that the two-storey ‘a’ has a higher visibility than the one-storey ‘a’, that the upper part of the bowl of the two-storey ‘a’ should be round and not diagonal, that a long tail on ‘j’, ‘l’ and ‘t’ enhance visibility at distance, and that the same goes for a serif on top of the stem of ‘i’. Furthermore, all fonts have an alternative **Stylistic Set** feature of the unfamiliar letter variations that delivered a good performance in the investigation.

*Sofie Beier*
*Miles Albert Tinker*

**Legibility, then, is concerned with perceiving letters and words and with the reading of continuous textual material.**

The shapes
of letters must be 
**DISCRIMINATED**  
the characteristic 
**word forms**  
PERCEIVED  
and continuous text read  
**ACCURATELY**  
rapidly  
**EASILY**  
and with understanding

*From the book 'Legibility of Print' by Miles A. Tinker*
TINKER EXTRA LIGHT ROMAN
ABCDEFHAbcdedefgh
01234567 ?!@&%#

TINKER EXTRA LIGHT ITALIC
ABCDEFHAbcdedefgh
01234567 ?!@&%#

TINKER LIGHT ROMAN
ABCDEFHAbcdedefgh
01234567 ?!@&%#

TINKER LIGHT ITALIC
ABCDEFHAbcdedefgh
01234567 ?!@&%#

TINKER REGULAR ROMAN
ABCDEFHAbcdedefgh
01234567 ?!@&%#

TINKER REGULAR ITALIC
ABCDEFHAbcdedefgh
01234567 ?!@&%#

TINKER MEDIUM ROMAN
ABCDEFHAbcdedefgh
01234567 ?!@&%#

TINKER MEDIUM ITALIC
ABCDEFHAbcdedefgh
01234567 ?!@&%#

TINKER BOLD ROMAN
ABCDEFHAbcdedefgh
01234567 ?!@&%#

TINKER BOLD ITALIC
ABCDEFHAbcdedefgh
01234567 ?!@&%#
...when one arrangement is read significantly faster, legibility must be a factor of importance.

Tinker extra light roman + bold 28/40 point
The experimental studies found that the most visible ‘l’ had a long tail. To accommodate this and avoid gaps inside the words, an ‘l’ placed at the end of a word is broader than an ‘l’ placed at the beginning or middle of a word.

MILES ALBERT TINKER

Tinker extra light roman 18/26 point
...when one arrangement is read significantly faster, legibility must be a factor of importance.

Tinker extra light roman small caps 18/26 point
...WHEN ONE ARRANGEMENT IS READ SIGNIFICANTLY FASTER, LEGIBILITY MUST BE A FACTOR OF IMPORTANCE.

Tinker extra light italic 18/26 point
...when one arrangement is read significantly faster, legibility must be a factor of importance.

Tinker extra light italic small caps 18/26 point
...WHEN ONE ARRANGEMENT IS READ SIGNIFICANTLY FASTER, LEGIBILITY MUST BE A FACTOR OF IMPORTANCE.
...when one arrangement is read significantly faster, legibility must be a factor of importance.

...WHEN ONE ARRANGEMENT IS READ SIGNIFICANTLY FASTER, LEGIBILITY MUST BE A FACTOR OF IMPORTANCE.

...when one arrangement is read significantly faster, legibility must be a factor of importance.

...WHEN ONE ARRANGEMENT IS READ SIGNIFICANTLY FASTER, LEGIBILITY MUST BE A FACTOR OF IMPORTANCE.

...when one arrangement is read significantly faster, legibility must be a factor of importance.

...WHEN ONE ARRANGEMENT IS READ SIGNIFICANTLY FASTER, LEGIBILITY MUST BE A FACTOR OF IMPORTANCE.

...when one arrangement is read significantly faster, legibility must be a factor of importance.

...WHEN ONE ARRANGEMENT IS READ SIGNIFICANTLY FASTER, LEGIBILITY MUST BE A FACTOR OF IMPORTANCE.
when one arrangement is read significantly faster, legibility must be a factor of importance.

...when one arrangement is read significantly faster, legibility must be a factor of importance.

...when one arrangement is read significantly faster, legibility must be a factor of importance.

...when one arrangement is read significantly faster, legibility must be a factor of importance.

...when one arrangement is read significantly faster, legibility must be a factor of importance.

...when one arrangement is read significantly faster, legibility must be a factor of importance.

...when one arrangement is read significantly faster, legibility must be a factor of importance.

...when one arrangement is read significantly faster, legibility must be a factor of importance.
legibility of isolated letters is also quite a different thing from legibility or readability of meaningful text.
Legibility of isolated letters is also quite a different thing from legibility or readability of meaningful text. Legibility of isolated letters is also quite a different thing from legibility of isolated letters. Legibility of isolated letters is also quite different. Legibility of isolated letters is also quite a different thing. Legibility of isolated letters is also quite different.

Tinker & Paterson 1949
In the final analysis, one wants to know what typographical factors foster ease and speed of reading. Optimal legibility of print, therefore, is achieved by a typographical arrangement in which shape of letters and other symbols, characteristic word forms, and all other typographical factors such as type size, line width, leading, etc., are coordinated to produce comfortable vision and easy and rapid reading with comprehension. In other words, legibility deals with the coordination of those typographical factors inherent in letters and other symbols, words, and connected textual material which affect ease and speed of reading.
In the final analysis, one wants to know what typographical factors foster ease and speed of reading. Optimal legibility of print, therefore, is achieved by a typographical arrangement in which shape of letters and other symbols, characteristic word forms, and all other typographical factors such as type size, line width, leading, etc., are coordinated to produce comfortable vision and easy and rapid reading with comprehension. In other words, legibility deals with the coordination of those typographical factors inherent in letters and other symbols, words, and connected textual material which affect ease and speed of reading.

In the final analysis, one wants to know what typographical factors foster ease and speed of reading. Optimal legibility of print, therefore, is achieved by a typographical arrangement in which shape of letters and other symbols, characteristic word forms, and all other typographical factors such as type size, line width, leading, etc., are coordinated to produce comfortable vision and easy and rapid reading with comprehension. In other words, legibility deals with the coordination of those typographical factors inherent in letters and other symbols, words, and connected textual material which affect ease and speed of reading.

In the final analysis, one wants to know what typographical factors foster ease and speed of reading. Optimal legibility of print, therefore, is achieved by a typographical arrangement in which shape of letters and other symbols, characteristic word forms, and all other typographical factors such as type size, line width, leading, etc., are coordinated to produce comfortable vision and easy and rapid reading with comprehension. In other words, legibility deals with the coordination of those typographical factors inherent in letters and other symbols, words, and connected textual material which affect ease and speed of reading.

In the final analysis, one wants to know what typographical factors foster ease and speed of reading. Optimal legibility of print, therefore, is achieved by a typographical arrangement in which shape of letters and other symbols, characteristic word forms, and all other typographical factors such as type size, line width, leading, etc., are coordinated to produce comfortable vision and easy and rapid reading with comprehension. In other words, legibility deals with the coordination of those typographical factors inherent in letters and other symbols, words, and connected textual material which affect ease and speed of reading.
Other things being equal, the greater the enclosed white space of a letter, the greater the legibility.

The factor which make it possible to read one style of print at a greater distance than another may not be the same as those which lead to the reading of one type face faster than another under ordinary reading conditions.

Miles A. Tinker 1964

Tinker & Webster 1935
Ovink

APPENDIX C
The typeface Ovink, named after the legibility researcher Gerrit Willem Ovink, is a Sans Serif typeface inspired by the Danish signage tradition. An early version of the typeface was subjected to experimental legibility investigations of distance and time threshold methods. With the typeface Ovink, participants were exposed to different variations of the most frequently misread lowercase letters. The findings, which were implemented in the final designs, demonstrated that the two-storey ‘a’ has a higher visibility than the one-storey ‘a’, that a diagonal spine of the ‘s’ has a lower visibility than a round spine, that the ‘i’ is more visible with a serif on top, that the ‘j’, ‘l’ and ‘t’ all benefit from a long tail, and that the ‘e’ at distance viewing is more visible with an open aperture. All fonts have an alternative Stylistic Set feature of the unfamiliar letter variations that delivered a good performance in the investigation.

Several experimental legibility investigations have found wider characters to be more legible at distance than narrow characters. Based on these findings, the typeface Ovink contains fonts of two widths, one Normal width to save horizontal space, and one Expanded width for maximum distance legibility. All weights are supplied in both positive and negative versions.

Sofie Beier
First, the legibility of isolated characters has been neglected hitherto, while it is at the same time highly important for the science and art.
of printing

and advertising, and experimentally

more accessible

THAN THE

legibility

of continuous reading-matter

From the top: OvinkExpandedRegularItalicBL, OvinkExpandedRegularBoldBL, OvinkExpandedRegularItalicBL, OvinkExpandedRegularRomanWH, OvinkExpandedBoldItalicWH, OvinkExpandedRegularRomanWH, OvinkexpandedBoldRomanBL, OvinkExpandedRegularItalicWH

From the book Legibility, Atmosphere-value, and Forms of Printing Types, by G.W. Ovink 1938
<table>
<thead>
<tr>
<th>Character set of Ovink Normal Regular Roman WH</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCDEFGHABCDEFg 0123456 ?!@&amp;%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ovink Normal Regular Italic WH</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCDEFGHABCDEFg 0123456 ?!@&amp;%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ovink Normal Bold Roman WH</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCDEFGHABCDEFg 0123456 ?!@&amp;%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ovink Normal Regular Roman BL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCDEFGHABCDEFg 0123456 ?!@&amp;%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ovink Normal Regular Italic BL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCDEFGHABCDEFg 0123456 ?!@&amp;%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ovink Normal Bold Roman BL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCDEFGHABCDEFg 0123456 ?!@&amp;%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ovink Normal Bold Italic BL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCDEFGHABCDEFg 0123456 ?!@&amp;%</td>
</tr>
<tr>
<td>Character set of OvinkExpandedRegularRomanWH</td>
</tr>
<tr>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz änšt</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
The subjective facility, or feeling of ease, has not got the place in the literature on our subject that it undoubtedly deserves.
The subjective facility, or feeling of ease, has not got the place in the literature on our subject that it undoubtedly de
The subjective facility, or feeling of ease, has not got the place in the literature on our su
The subjective facility, or feeling of ease, has not got the place in the literature
The subjective facility, or feeling of ease, has not got the place in
The subjective facility, or feeling of ease, has not got the p
The subjective facility, or feeling of ease, has not got t
The subjective facility, or feeling of ease, has no
The subjective facility, or feeling of ease, has
The subjective facility, or feeling of ease, has e
The subjective facility, or feeling of e
The subjective facility, or feeling o
The subjective facility, or feeling
The subjective facility, or feelin
The subjective facility, or feeli
The subjective facility, or fee
The subjective facility, or fe
The subjective facility, or f
The subjective facility, or f
Only a few investigators used isolated letters: others however, used words or lines. In both cases no different forms of the same character were compared, but rather the relative possibility of interchange was searched between the various characters of the alphabet at one and the same typeface.
Only a few investigators used isolated letters: others however, used words or lines. In both cases no different forms of the same character were compared, but rather the relative possibility of interchange was searched between the various characters of the alphabet at one and the same typeface.
Only a few investigators used isolated letters: others however, used words or lines. In both cases no different forms of the same character were compared, but rather the relative possibility of interchange was searched between the various characters of the alphabet at one and the same typeface.
Only a few investigators used isolated letters: others however, used words or lines. In both cases no different forms of the same character were compared, but rather the relative possibility of interchange was searched between the various characters of the alphabet at one and the same typeface.
THE STUDY OF THE SHAPES OF SINGLE CHARACTERS ONLY,

means an arbitrary limitation of possible causes; it has to be done side by side of research of assimilative

...it is often far more important for a publisher or advertiser to know whether the reading public dislike a type or not, than to know the actual differences in speed, which are small anyhow.

THE STUDY OF THE SHAPES OF SINGLE CHARACTERS only, means an arbitrary limitation of possible causes it has to be done side by side of research of ASSIMILATIVE READING

G.W. Ovink 1938
Pyke

APPENDIX D
The typeface Pyke, named after the legibility researcher Richard Lionel Pyke, is a serif typeface inspired by the work of Giambattista Bodoni. An early version of the typeface was subjected to experimental legibility investigations of distance and time threshold methods. With the typeface Pyke, participants were exposed to different variations of the most frequently misread lowercase letters. The findings, which were implemented in the final designs, demonstrated that in a short exposure the italic style ‘l’ with a tail has a higher visibility than the ‘l’ with serifs at the bottom, that the italic style descending ‘f’ is more visible at distance than the Roman style ‘f’, that a closed aperture of the ‘e’ lowers visibility in a short exposure, and that the same goes for the closed aperture of the ‘c’ at distance. Furthermore, all fonts have an alternative Stylistic Set feature of the unfamiliar letter variations that delivered a good performance in the investigation.

Following the tradition of Bodoni, fonts of the typeface family Pyke are size specific, including versions for Micro, Text, and Display.

Sofie Beier
The hypothesis is here put forward that extremely large typographical differences must be present.
before it is possible

**PYKE DISPLAY REGULAR ROMAN**

to say that there is

**PYKE DISPLAY BOLD ITALIC**

any difference in the

**PYKE TEXT BOLD ROMAN**

**OBJECTIVE LEGIBILITY**

**PYKE DISPLAY BOLD ROMAN**

of

**PYKE DISPLAY REGULAR ITALIC**

**TYPES**

**PYKE DISPLAY BOLD ROMAN**

## Pyke Character Set

<table>
<thead>
<tr>
<th>Micro</th>
<th>Text</th>
<th>Display</th>
</tr>
</thead>
<tbody>
<tr>
<td>A B C D E F G H a b c d e f g h</td>
<td>0 1 2 3 4 5 6 7</td>
<td>! @ &amp; # %</td>
</tr>
<tr>
<td>a b c d e f g h a b c d e f g h</td>
<td>0 1 2 3 4 5 6 7</td>
<td>! @ &amp; # %</td>
</tr>
<tr>
<td>a b c d e f g h a b c d e f g h</td>
<td>0 1 2 3 4 5 6 7</td>
<td>! @ &amp; # %</td>
</tr>
<tr>
<td>a b c d e f g h a b c d e f g h</td>
<td>0 1 2 3 4 5 6 7</td>
<td>! @ &amp; # %</td>
</tr>
<tr>
<td>a b c d e f g h a b c d e f g h</td>
<td>0 1 2 3 4 5 6 7</td>
<td>! @ &amp; # %</td>
</tr>
</tbody>
</table>

**Pyke Micro Regular Roman**

| ABCDEFGHabcdefgh | 01234567 !@&##%

**Pyke Micro Regular Italic**

| ABCDEFGHabcdefgh | 01234567 !@&##%

**Pyke Micro Bold Roman**

| ABCDEFGHabcdefgh | 01234567 !@&##%

**Pyke Micro Bold Italic**

| ABCDEFGHabcdefgh | 01234567 !@&##%

**Pyke Text Regular Roman**

| ABCDEFGHabcdefgh | 01234567 !@&##%

**Pyke Text Regular Italic**

| ABCDEFGHabcdefgh | 01234567 !@&##%

**Pyke Text Bold Roman**

| ABCDEFGHabcdefgh | 01234567 !@&##%

**Pyke Text Bold Italic**

| ABCDEFGHabcdefgh | 01234567 !@&##%

**Pyke Display Regular Roman**

| ABCDEFGHabcdefgh | 01234567 !@&##%

**Pyke Display Regular Italic**

| ABCDEFGHabcdefgh | 01234567 !@&##%

**Pyke Display Bold Roman**

| ABCDEFGHabcdefgh | 01234567 !@&##%

**Pyke Display Bold Italic**

| ABCDEFGHabcdefgh | 01234567 !@&##%
Legibility must be distinguished from terms like recognizability, perceptibility, &c.
Legibility must be distinguished from terms like recognizability, perceptibility, &c.
Legibility must be distinguished from terms like recognizability, perceptibility, &c.

Legibility must be distinguished from terms like recognizability, perceptibility, &c.

Legibility must be distinguished from terms like recognizability, perceptibility, &c.

Legibility must be distinguished from terms like recognizability, perceptibility, &c.

Legibility must be distinguished from terms like recognizability, perceptibility, &c.

Legibility must be distinguished from terms like recognizability, perceptibility, &c.
Whether reading at maximum speed distorts the true influence of objective legibility is not known, nor apparently is the effect of maximum speed on errors.
Whether reading at maximum speed distorts the true influence of objective legibility is not known, nor apparently is the effect of maximum speed on errors.
The problem of legibility seemed simple at the outset; it is in fact complex and elusive, and many questions have been raised which have not been answered. Things being what they are, this is all to the good. For the practical, typographical problems of legibility are still far—further than many have believed—from scientific solution.

Such solution is not possible until the manifold problems of principle, method, and technique have been acknowledged to exist and themselves solved. For this purpose are needed fewer opinions and more facts.
The problem of legibility seemed simple at the outset; it is in fact complex and elusive, and many questions have been raised which have not been answered. Things being what they are, this is all to the good. For the practical, typographical problems of legibility are still far—further than many have believed—from scientific solution. Such solution is not possible until the manifold problems of principle, method, and technique have been acknowledged to exist and themselves solved.

For this purpose are needed fewer opinions and more facts.
ILLEGIBILITY

Manifests Itself

Not by mistakes so much as by slow reading

On the other hand, these facts practically do not apply to nonsense or a foreign language, for which errors are a better index. If this is true, speed is not reliable here; it may vary directly and not inversely with accuracy. If this, too, is true, an increase in errors alone might mean merely

AGAIN, IN SENSE,

to measure by speed is more realistic

* not to say logical than by errors for in normal life one reads just so fast that one does

NOT MAKE ERRORS

Testing for legibility can best be done

by

ACTUAL READING

R.L. Pyke 1926
DEFINITIONS

&

BIBLIOGRAPHY
Glossary

**Agate**: Fonts designed for very small sizes.

**Artefact**: An object that has been intentionally made or produced for a certain purpose.

**Blackletter**: The term originated in the US, where ‘Gothic’ (the European term for the style) was the name of Sans Serif faces.

**Body**: The block from which the glyph rises in metal and wood type.

**Cones**: The cone cells are concentrated in the fovea of the eye, and gradually become sparser towards the periphery of the retina. Cones operate in bright light; they specialize in processing details and have a high visual acuity.

**Contrast**: The relationship between the thin and the thick parts of the stroke.

**Counterpunch**: There are two ways of creating the counters of letters in punchcutting: one is by digging the counter out with a graver; the other is by striking a counterpunch into the punch.

**Crowding**: Refers to the phenomenon of letters placed next to each other visually merging, causing individual shapes that are difficult to tell apart.

**Didone**: A typeface category in the Vox-ATypI classification and the British Standards. Derived from the names of Didot and Bodoni, the style is described as follows: Typefaces having an abrupt contrast between thin and thick strokes; the axis of the curves is vertical; the serifs of the ascenders of the lower case are horizontal; there are often no brackets to the serif.

**Display typefaces**: Designed for headings, not to be mistaken for signage typefaces.
Distance methodology: Measuring visual acuity at a distance.

Extenders: Referring to both ascending and descending parts of letters.

Familiarity: The collective influence of previous exposure and the level of common letter features.

Features [of letters]: The individual details, such as the eye of the 'e' or the tail of the 't'.

Font: In metal and word type, this refers to the complete character set of a single size and style of a particular typeface. In digital type, it refers to a single weight of a particular typeface.

Foveal: The area around the fixation point covering about 2 degrees.

Fitting: Inter-character spacing.

Fixation in reading: The relatively stable movement between saccades, where the eye rests on the text.

Geralde: A typeface category in the Vox-ATypI classification and the British Standards. Derived from the names of Garamond and Aldus Manutius, the style is described as follows: Typefaces in which the axis of the curves is inclined to the left; there is generally a greater contrast in the relative thickness of the strokes than in Humanist designs; the serifs are bracketed; the bar of the lowercase e is horizontal; the serifs of the ascenders in the lower case are oblique.

Glyph: In digital type design the term refers to a single, specific version of a character that appears in a font.

Gothic: A general term used to describe the heavy angular scripts were black dominates white on the page. As subcategories we find the picket fence 'Textura' mostly associated with northern Europe, the more rounded 'Rontunda' from southern Europe, and 'Schwabacher' and 'Fraktur' for the German language.
Hand-compositing: Manual typesetting

**Humanist:** A typeface categories in the Vox-ATypI classification and the British Standards. The serif style is described as follows: Typefaces in which the cross stroke of the lower case ‘e’ is oblique; the axis of the curves is inclined to the left; there is no great contrast between thin and thick strokes; the serifs are bracketed; the serifs of the ascenders in the lower case are oblique. The sans serif style is described as follows: Lineale typefaces based on the proportions of inscriptionsal Roman capitals, rather than on early grotesque. They have some stroke contrast, with two-storey ‘a’ and ‘g’.

**Interaction effect:** In the familiarity investigation, it is the interaction between the font conditions and the exposure condition. Essentially that one font will improve more from exposure than will another font.

**Italic:** A font with characters such as ‘a’, ‘e’, ‘f’ and ‘g’ often being more cursive in style than what is seen in its Roman counterpart.

**Kerning:** The adjustment of the space between characters. For example an ‘A’ and a ‘T’ can be kerned more tightly, so they appear optically correct when viewed together in a word.

**Legibility:** Clarity of the individual letters influenced by familiarity.

**Lettering:** Drawing of letters where the brush or pen changes angle to achieve a specific style.

**Linotype line-caster machine:** Matrices are first put together in a line. When the line is complete it is cast as a single piece of metal. It is not possible to kern characters.

**Main effect:** Example in the familiarity investigation: the pre-test and the post-test data for each font are combined for comparison between the three font conditions.
Monotype compositing machine: In contrast to Linotype’s line-caster, each character is here cast as an individual piece of metal. Kerned letters are possible.

Nonsense words: Unpronounceable words.

Old Style: A collective reference to the Humanist and Geralde serif faces of the typeface classification system of Vox-ATypI and the British Standards.

One-storey: Lowercases ‘a’ and ‘g’ created out of one loop instead of two.

Parafoveal: Outside the foveal area is the parafoveal area, which contains about 4 degrees to the left and to the right.

Peripheral: Everything outside the parafoveal area is the peripheral.

PLR Model: Short for the Parallel Letter Recognition Model

Pseudowords: Pronounceable nonexistent words.

Rapid serial visual presentations (RSVP): Also known as ‘Rapid Serial Word Presentation’, is a method of showing text where each piece of the text is displayed briefly in the same location and in sequential order.

Repeated measure: The same subjects are used in both the pre-test and the post-test of the familiarity investigation.

Readability: Ease of reading in running text.

Rods: The rod cells are concentrated at the outer edges of the retina and are used in peripheral vision. Rods are specialized for detecting movement and small differences in brightness. They have a low visual acuity.

Roman: The upright version, in relation to the Italic version.
Saccade: When we read, the eye completes a series of jerky movements interrupted by fixations.

Sans Serif: Typefaces without serifs.

Search task methodology: The task is to locate a specific letter or word in a given text.

Short exposure methodology: The task is to identify an object after a short exposure.

Skeleton: The basic shape of a letter isolated from weight, contrast, width, proportions and stroke.

Slab Serif: A typeface category in the Vox-ATypI classification and the British Standards. The style is described as follows: Typefaces with heavy, square-ended serifs, with or without brackets.

Speed of reading methodology: Measuring the time it takes a participants to read a given text.

Stereotyping: A printing plate cast in a mold and made from composed type or an original plate. Developed in the late 18th century.

Style: The collective influence of serifs, contrast, stroke, proportion and weight.

Traditional legibility research: Typeface comparison studies, aiming at locating the most legible sample within the tested material.

Transitional: A typeface category in the Vox-ATypI classification and the British Standards. The style is described as follows: Typefaces in which the axis of the curves is vertical or inclined slightly to the left; the serifs are bracketed, and those of the ascenders in the lower case are oblique.

Typeface: The look and design of a collection of fonts.
Typeface family: A series of related fonts including roman, italic, regular and bold weights, etc.

Two-storey: Lowercases ‘a’ and ‘g’ created out of two loops instead of one.

Visibility: Clarity of shapes isolated from the influence of familiarity.

Visual angle: The relationship between the size of an object and the distance to the eye.

Weight: The visual darkness or lightness of the strokes in a character.
Bibliography

A


Aicher, O. & Rommen, J. (1988) Typographie, Ernst & Sohn Verlag


André, J. & Girou, D. (1999) Father Truchet, the typographic point, the Romain du roi, and tilings, TUGboat, 20(1), 8-14


Aynsley, J. (2000) Graphic Design in Germany 1890-1945, Thames & Hudson


Baines, P. (1999/00) ‘Letterforms surround us’, Point: art and design research journal, 8, 5-13


Baudin, F. (1967a) ‘Miles A. Tinker, Bases for effective reading’, Journal of Typographic Research, 1(2), 204-207


Beaujon, P. [pen-name of Beatrice Warde] (1933) ‘On the Choice of Type Faces’, The Monotype Recorder, 32. 5-10


Belknap, J. (1993) ‘If the face fits’. Eye, 3(11), 56-63


Berry, J.D. (2004) Now read this: The Microsoft ClearType Font Collection, J.D. Berry & J. Hudson, (eds), Microsoft Corporation

Berry, J.D. (2006a) Dot-font: Talking About Fonts, New York: Mark Batty Publisher


British Standards Institution (1967a) Specification for conventional typographical character for legibility tests (ISO test character), London: BSI

British Standards Institution (1967b) BS 2961:1967, Typeface Nomenclature and Classification, London: BSI


Bodoni, G. (1818) Manuale Tipografico, Parma: Presso la Vedona


C


Cabarga, L. (2004b) Logo, Font & Lettering Bible: A comprehensive guide to the design, construction and usage of alphabets, letters and symbols, Davis & Charles


Chapman, J. (1978) ‘Re-thinking research into visual communication’, Icographic, 13, 28-32


Chaytor, H.J. (1945) From Script to Print, Cambridge: At the University Press


Crosby/Fletcher/Forbes (1970) A Sign Systems Manual, Praeger Publisher


Davis, A. (1951) ‘Type in Advertising’, Raithby, Lawrence & Co


Dreyfus, J. (1951) ‘Baskerville’s Methods of Printing’, Signature: new series, 12, 43-51


The Art of the Type Specimen in the Twentieth Century, New York: The 
Typophiles

Dreyfus, J. (1994) Into Print: Selected Writings on Printing History, Typog-
raphy and Book Production, London: British Library


York: Dover Publications


Dunn–Rankin, P. (1968) ‘The Similarity of Lower-Case Letters of the 
English Alphabet’, Journal of Verbal Learning and Verbal Behaviour, 7(6), 
990-995

Dunn–Rankin, P., Leton, D.A. Shelton, V.F. (1968) ‘Congruency factors 
related to visual confusion of English letters’, Perceptual & Motor Skills, 
26, 659-666

Dwiggins, W.A. (1928) Layout in Advertising, New York: Harper and 
Brothers Publishing

Dwiggins, W.A. (1940) WAD to RR: a letter about designing Type, Harvard 
College Library

Dwiggins, W.A. (1947) Mss. by WAD: Being a collection of the writings 
of Dwiggins on various subjects Some critical, some philosophical, some 
whimsical, New York: The Typophiles

and line length on the effectiveness of reading from screen’. International 
Journal of Human-Computer Studies, 54, 585-612

Dyson, M.C. & Kipping, G.J. (1997) ‘The Legibility of Screen Formats: 
are Three Columns Better than One?’. Computers & Graphics, 21(6), 
703-712

Hyphen, 2(1), 5-13

Dyson, M.C. (1999/00) ‘Typography’s contribution to the design of 
human-computer interfaces’, POINT, 8, 22-25

screen’, Behaviour & Information Technology, 23(6), 377-393

for guidance?’. Typo, 13, 30-37
E


Ellegaard Frederiksen, E. (1955) Bogens Funktion og Æstetik, Copenhagen: Christian Ejler’s Forlag

Ellegaard Frederiksen, E. (1965) Knud V. Engelhardt, Arkitektens Forlag


Eurographic Press Interview (1962) ‘Designer’s profile: Adrian Frutiger’, Print in Britain, 9(9), 258-262

Ewald, F. (1930) ‘The Officina Bodoni’, The Fleuron, 7, 121-131

F


G


Gaultney, V. (2001) Balancing typeface legibility and economy, unpublished essay, The University of Reading, Department ofTypography & Graphic Communication


**Goudy, F.W.** (1940) *Typologia: studies in type design & type making, with comments on the invention of typography, the first types, legibility, and fine printing*, Berkeley: University of California Press.


Haber, L.R., Haber, R.N. & Furlin, K.R. (1983) 'Word length and word shape as sources of information in reading', Reading Research Quarterly, 18(2), 165-189


Haddon Craftsmen (1937) A Primer of Types: Giambattista Bodoni 1740-1813, Camden

Haley, A. (1987) 'Giambattista Bodoni', U&lc, 14(3), 14-17


Harling, R. (1946) 'The Early Alphabets of Eric Gill', Alphabet and Image, 3, 61-67


Handover, M. (1960) 'Palette for printers', Motif, 5, 94-95


Harris, J. (1973) 'Confusions in letter recognition', Professional Printer, 17(2). 29-34


HCM/ Graphic System (1981) ‘Demos and Praxis’ (advert), U&lc, 7(4), 60


Helvetica (2006), Directed by Gary Hustwit, Plexifilm [film: DVD]


Hutt, A. (1946) 'The Gothic Title-Piece and the English Newspaper’. Alphabet and Image, 3. 3-19

Hutt, A. (1962) 'Newspaper text in the sixties’, Print in Britain, 10(5). 24-25


Hutt, A. (1972) Fournier: the complete typographer, London: Frederick Muller LTD
I

J


Jones, A. (2001) 'Do we need more or fewer typefaces', in: D. Jury, (ed), Typographic Writing, ISTD, 21-24


K


Karow, P. (1994b) Digital Typefaces: Description and Formats, Berlin Herdelberg: Springer-Verlag


Kennedy, R. (2003) 'Sans Serif', Type & Typography, West New York: Mark Batty, Publisher, 319-321


Kindersley, D. (1960) 'Motorway sign lettering', Traffic Engineering & Control, 2(8), 463-465

Kinneir, J. (1978) 'The practical and graphic problems of road sign design', in: R. Easterby & H. Zwaga, (eds), Information design: the design and evaluation of signs and technical information, Chichester: John Wiley and Sons Ltd., 341-58


Kinross, R. (1986) ‘What is a typeface?’, *Baseline*, 7, 14-18


Krimer, J.V. (1952) On Designing and Devising Type, New York: The Typophiles


Kuennapas, T. & Janson, A. (1969) 'Multidimensional similarity of letters', Perceptual & Motor Skills, 28, 3-12

L


Lane, B. (1966) 'Legibility depends on the reader', Print in Britain, 14(8), 12

Lane, J.A. (2005) 'Claude Garamont and his Roman Type, Garamond Premier Pro: A Contemporary Adaptation, Adobe, 5-13


Latham, J. (1979) 'New technology and good typography are not incompatible', Monotype Recorder, new ser., 1, December, 25-26


Legros, L.A. (1922) A Note on the Legibility of Printed Matter: Prepared for the information of the Committee on Type Faces, Joint Author (with John Cameron Grant) of the typographical Printing-Surfaces, London: Printed & Published by his Majesty's Stationary Office.


Lewis, J. (1962) Printed Ephemera, W.S. Cowell LTD.


Lund, O. (1999) Knowledge Construction in Typography: The case of legibility research and the legibility of sans serif typefaces. Thesis (PhD), The University of Reading, Department of Typography & Graphic Communication


Mardersteig, G. (1968b) On G.B. Bodoni’s Type Faces, Verona


Mergenthaler Linotype Company (1935) The Legibility of Type, Brooklyn N.Y.

Middendorp, J. (2004) Dutch Type, Rotterdam: 010 publishers


Monotype Recorder (1933) ‘An Account of The L·N·E·R Type Standardization’, The Monotype Recorder, 32(4), 7-11


Morison, S. (1924) ‘Towards an Ideal Type’, The Fleuron, 2, 57-75


Morison, S. (1949b) Four Centuries of Fine Printing, London: Ernest Benn Limited


Moxon, J. (1683) *Mechanick Exercises or the Doctrine of Handy-works*, London: Joseph Moxon


N


Novarese, A. (1963) 'Designing Eurostile', Print in Britain, 10(11), 18-19


Ovink, G.W. (1938) Legibility, Atmosphere-value, and Forms of Printing Types, Leiden


Poulton, E.C. (1972) 'Size, style, and vertical spacing in the legibility of small typefaces', Journal of Applied Psychology, 56(2). 156-161


Poulton, E.C. (1972) 'Size, style, and vertical spacing in the legibility of small typefaces', Journal of Applied Psychology, 56(2). 156-161


Reynolds, L. (1979b) 'Legibility studies: their relevance to present day documentation methods', *Journal of Documentation*, 35(4), 307-340

Reynolds, L. (1979c) 'The Graphic Information research Unit: background and recent research', *Visible Language*, 1(4), 428-448


Rogers, B. (1949) The Centaur Types, Chicago: October House


Schnell, T. (1999) *Legibility Threshold Contrast of Uppercase Text seen against a Dark Background* [online], Ohio University, available from: www.ent.ohiou.edu/ce/orite [Accessed 20 August 2007]


Schwartz, C. (2006a) 'Back with a Flourish', *Eye*, 16, 44-49


Smith, N.B. (1928) 'Matching ability as a factor in first grade reading’ *Journal of Educational Psychology*, 19, 560-571


Spencer, H., Reynolds L. & Coe, B. (1973b) A comparison of the effectiveness of selected typographic variations, Readability of Print Research Unit, Royal College of Art


Spencer, H. (1968a) The visible word, London: The Royal College of Art


Spiekermann, E. (1986) ‘Post Mortem, or: How I once designed a typeface for Europe’s biggest company’, Baseline, 7, 6-8


Soar, R.S. (1955) ‘Stroke Width, Illumination Level, and Figure-Ground Contrast in Numerical Visibility’, *Journal of Applied Psychology*, 39(6), 429-432


Tankard, J. (2004) TypeBookOne, Jeremy Tankard Typography Ltd


Tinker, M.A. & Paterson D.G. (1941) ‘Eye Movements in reading a Modern Type Face and old English’, American Journal of Psychology. 54, 113-114

Tinker, M.A. & Paterson, D.G. (1949) ‘What is Legibility’, Print. 6(2), 61


Tinker, M.A. (1928) ‘The relative legibility of the letters, the digits, and of certain mathematical signs’, Journal of General Psychology. 1, 472-496


Tinker, M.A. (1939) ‘Speed versus Comprehension in Reading as Affected by Level of Difficulty’, Journal of Educational Psychology. 30, 81-94

Tinker, M.A. (1940) ‘Dr Robinson on speed versus comprehension in reading: a discussion’, Journal of Educational Psychology. 31, 559-560


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


Tracy, W. (1952) ‘Newspaper Typography’, Typographica, 7, 5-19


U


Unger, G. (2007a) While You’re Reading, New York: Mark Batty Publisher


Uttal, W.R. (1969) ‘Masking of alphabetic character recognition by dynamic visual noise (DVN)’, Perception and Psychophysics, 6, 121-128


Waller, R. (1987) *The typographic contribution to language: towards a model of typographic genres and their underlying structures*, Thesis (PhD), Department of Typography and Graphic Communication, University of Reading


Walters, J.L. & Oliver, A. (2001) 'Gerard Unger'. *Eye*, 10(40), 10-16


Warde, B. (1927) 'The Baskerville Types a Critique'. *The Monotype Recorder*, 26(221) sept-oct. 3-29

Warde, B. (1936) *Printing Should be Invisible*, Middlesex: at the Chiswick Polytechnic
Warde, B. (1954) 'Typography means more than layout', Bowater Papers, 3. 42-46


Warde, B. (1956) 'New Light on typography legibility', Penrose Annual, 50. 51-55


Weidemann, K. (1985) 'Biblica — designing a new typeface for the Bible', Baseline, 6, 7-11


Wyse, H.T. (1911) Modern Type Display: And the use of Type Ornament, Henry T. Wyse
Y


Z


