

# John Benjamins Publishing Company



This is a contribution from *Information Design Journal 20:1*

© 2013. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: [www.copyright.com](http://www.copyright.com)).

Please contact [rights@benjamins.nl](mailto:rights@benjamins.nl) or consult our website: [www.benjamins.com](http://www.benjamins.com)

Tables of Contents, abstracts and guidelines are available at [www.benjamins.com](http://www.benjamins.com)

Sofie Beier and Kevin Larson

## How does typeface familiarity affect reading performance and reader preference?

**Keywords:** familiarity, legibility, readability, typeface design, fonts, experimental study, typographic research

Some typographers have proposed that typeface familiarity is defined by the amount of time that a reader has been exposed to a typeface design, while other typographers have proposed that familiarity is defined by the commonalities in letter shapes. These two hypotheses were tested by measuring the reading speed and preferences of participants. Participants were tested twice with common and uncommon letter shapes, once before and once after spending 20 minutes reading a story with the font. The results indicate that the exposure period has an effect on the speed of reading, but the uncommon letter shapes did not. Readers did not like the uncommon letter shapes. This has implications for the selection of type and the design of future typefaces.

### Introduction

Typographers believe that typeface familiarity plays an important role in the reading process, and use the theory of familiarity to make a number of arguments. Conservative typographers argue that readers are best supported by not making changes to the design of typefaces, while liberal typographers argue that we need to overcome the status quo to improve legibility. The conservative

approach is represented by type scholar Stanley Morison (1889–1967), who believes that a functional typeface should “be so good that only very few recognize its novelty” (Morison, 1930, p. 63); elsewhere Morison explains that for a typeface to be satisfactory, the essential form must “correspond with that handed down” (Morison, 1924, p. 59). In his later years, typographer Jan Tschichold (1902–1974) further supports this view by stating that “Our type is in fact an absolutely inflexible form, and offers no possibility for any but minute alterations if we still want to read with ease” (Tschichold, 1969, p. 53). A more liberal approach is advocated by type designer Eric Gill (1882–1940), who argues that, although we are used to material that is less legible than other material would be if it were equally familiar, we should still try to improve on the existing material (Gill, 1936). This view is later supported by William Addison Dwiggins, who notes, “If the reader is used to bad design, he must be led to accustom himself to better design” (Dwiggins, 1947, p. 40).

The question is then: should we follow the conservative view and not change the appearance of letterforms, or should we be more open to adjustments? In an attempt to find an answer, the focus of the present investigation is on two hypotheses: (1) that familiarity is based on common letter shapes, so typeface design must stay within strict parameters; and (2) that familiarity is based

on the amount of exposure to a particular typeface, so typeface design can change since it is just a matter of time before the new design becomes familiar as well.

### The prototype hypothesis

The first hypothesis argues that familiarity comes from a mental prototype for each letter that each of us has developed from a lifetime of exposure to common letter shapes. Type designer Adrian Frutiger's (1998) letter-

form matrix model argues that every character has a basic skeleton based on a collective memory of all the different character variations a person has ever encountered. The skeleton emerges when widely read typefaces are superimposed, so that the parts of the letterform that are shared across all typefaces are shown. Frutiger demonstrates this effect by using eight of what he defines as the most widely read typefaces: Garamond, Baskerville, Bodoni, Excelsior, Times, Palatino, Optima, and Helvetica. While Frutiger created his images using a

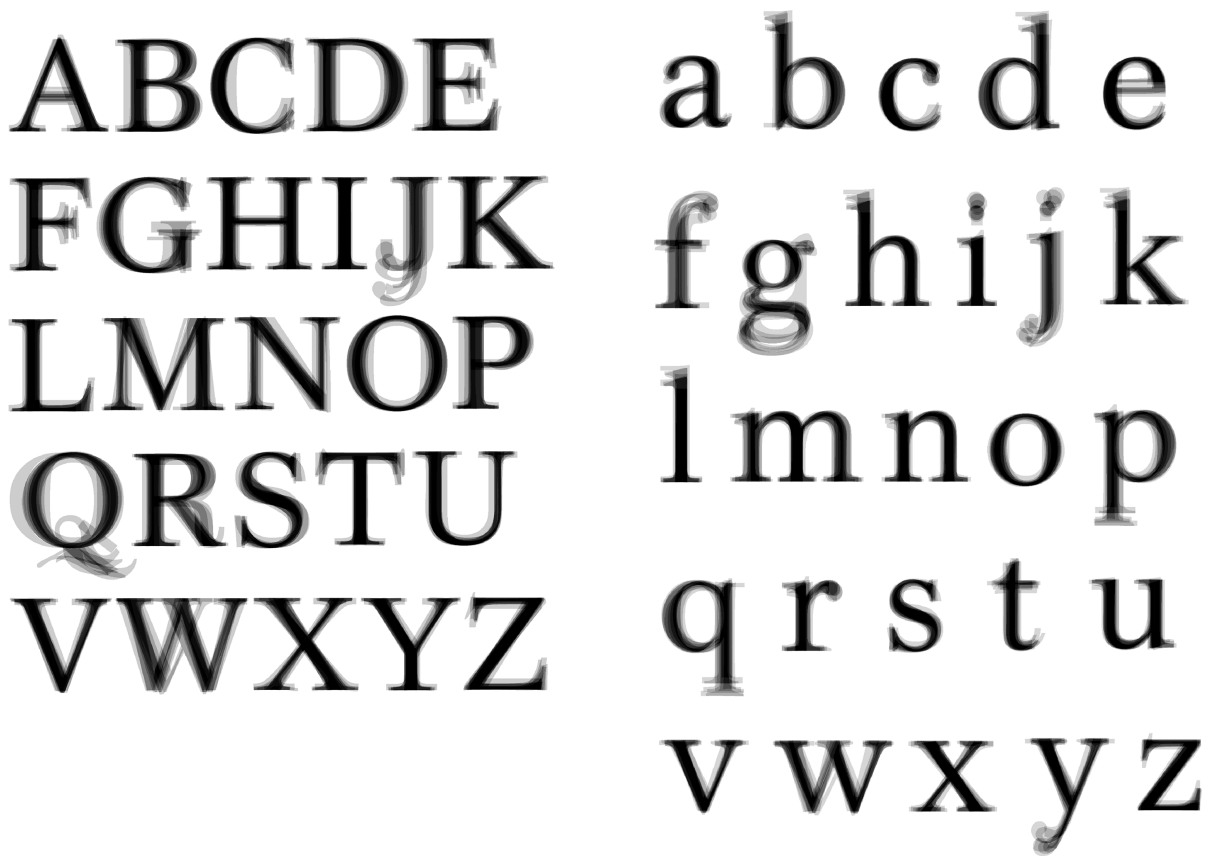


Figure 1. The complete a-z set of upper and lowercase letters for Adrian Frutiger's letterform matrix model.

phototypesetting machine, we recreated his experiment by superimposing digital fonts. Each font was scaled to a matching x-height for the lowercase and cap-height for the uppercase. The letters were shifted horizontally to the position that maximized overlap. The results can be seen in Figure 1. The white area is contained in no font, the lightest level of grey is contained in one font, darker grey is contained in multiple fonts, and the black area in all eight fonts.

Frutiger's letterform matrix provides a useful definition for prototypes. Some fonts have a very high degree of overlap with the matrix, while others have less. For example, the Garamond "a" has less overlap with the matrix because it is narrower than the other fonts and the bowl of the "a" is lower. For any letter in a particular font, it is possible to quantify the amount of overlap with the letterform matrix by correlating the pixels that are turned on in a particular font's letter with the darkness of the pixels in the letterform matrix. The correlation will be low if the target letter contains pixels that are light grey or white in the letterform matrix, or high if the target letter's pixels are completely contained in the black or dark grey area of the matrix.

The prototype model is supported by Langlois' research on the recognition of human faces. Langlois and Roggman (1990) found that composite faces, created by blending pictures of many individuals, were judged to be more attractive than individual faces. In another of Langlois' experiments mentioned by Dingfelder (2006), the researchers found that both toddlers and adults recognized attractive faces more quickly than ugly faces. Since we know that attractive faces are average faces, it suggests that prototypical faces are recognized faster than faces that are further from the prototypical face. If the Langlois discoveries can be transferred to typefaces, it suggests that common letter shapes are viewed as more beautiful, and also recognized more quickly, than uncommon letter shapes.

## The exposure hypothesis

The second hypothesis is that familiarity improves with exposure to a particular typeface design. This suggests that while new typefaces will be read inefficiently, readers will become efficient at reading a particular typeface by spending time reading with it. Advocating this view, the Emigre type designer Zuzana Licko famously states that "Readers read best what they read most" (Licko, 1990, p. 12). 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 made these comments at a time in her career when she was experimenting with some rather unusual looking typefaces. In the text she goes on to speculate whether in 200 years her own typefaces will be viewed as legible (Licko, 1990).

Research carried out by Zineddin, Garvey, Carlson, and Pietrucha (2003) supports the hypothesis that people read more efficiently with typefaces that they read frequently. Their study investigated the improvement on a distance threshold task after reading with three fonts with different familiarity levels in the letter shapes. 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 in performance for the fonts used in the exposure sessions, and established that exposure is part of familiarity.

## Font tuning

The exposure hypothesis is different from a related finding called “font tuning”. Font tuning is a quick-acting adjustment by the visual system. In an experimental investigation, Sanocki (1988) compared strings of letters from one font with strings of mixed letters from another font made out of letters from two different typefaces. One of the typefaces was a Sans Serif with rounded corners and a large x-height, while the other was a Slab Serif with square corners and a smaller x-height. The two typefaces did not share a baseline, which resulted in the mixed font jumping up and down visually. Setting out to identify the influence of familiarity 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).

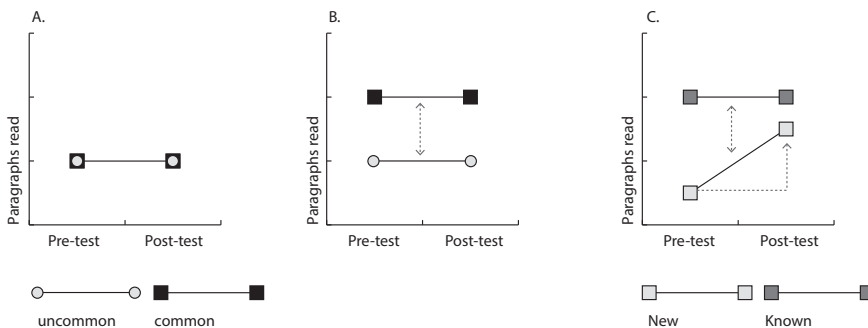
The finding that font tuning is a perceptual phenomenon has been extended in recent studies. Walker (2008) demonstrated font tuning in lexical decision-making where participants decided if a string of letters was a word or pseudo-word. Participants were faster at

completing this task if the preceding lexical decision was shown in the same font rather than a different one.

Gauthier, Wong, Hayward, and Cheung (2006) demonstrated that only experts in an orthography exhibit font tuning. They found that English readers viewing English, and Chinese readers viewing Chinese both exhibited font tuning, but that non-Chinese readers did not show evidence of font tuning when viewing Chinese. Additionally, they found that font tuning was stronger for changes in aspect ratio (x-height) than for the font manipulations of slant direction and outline fill.

## The investigation

The current investigation was designed to test directly the relative contributions of the prototype hypothesis and the exposure hypothesis on familiarity. Participants in this study read text with fonts that had common letter shapes, correlating strongly with Frutiger’s letterform matrix, and fonts that had uncommon shapes, correlating poorly with the letterform matrix. Additionally, the fonts were either known or unknown to the participants. Participants had their reading speed with each font measured both before and after a 20 minute exposure period. The prototype hypothesis predicts that the fonts with common letter shapes should be more familiar and



**Figure 2.** Illustration of possible outcomes. A: The null hypothesis will show no effect of familiarity condition and no effect of repeated measure. B: The finding expected for the prototype hypothesis is that there will be an effect of familiarity condition and no effect of repeated measure. C: The finding expected for the exposure hypothesis is that there will be an effect of repeated measure and no effect of familiarity.

quicker to read than fonts with uncommon letter shapes. The exposure hypothesis predicts that reading speed will increase after spending time reading with the font. Both hypotheses could be supported by finding that fonts with common letter shapes performed better than fonts of uncommon letter shapes both before and after the exposure period, but that both kinds of fonts were read faster after the exposure period.

## Method

Two studies were carried out with one font from each of the three familiarity conditions: known-common, unknown-common, and unknown-uncommon. Fonts of the known-common condition were well-known fonts of common letter shapes that all readers had been exposed to extensively before. Fonts of the unknown-common condition had common letter shapes but no prior exposure. Fonts of the unknown-uncommon had uncommon letter shapes and no prior exposure. A known-uncommon condition was not included, due to difficulties in identifying two typefaces that both had the same level of previous exposure as the known-common fonts and the same level of uncommon letter shapes as the new-uncommon fonts.

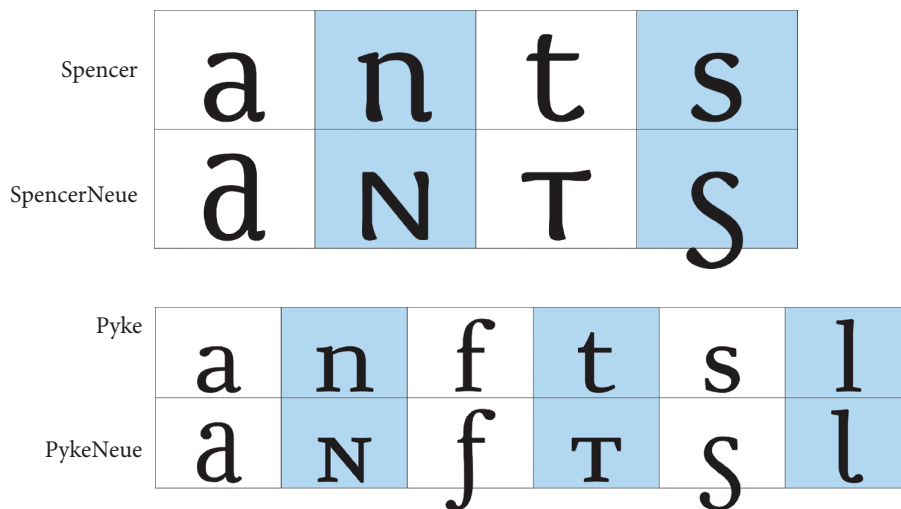
## Test material – controlling legibility

From the 1920s to the 1950s Miles A. Tinker (1964) ran an extensive program of research measuring characteristics about typefaces and text layout that resulted in faster reading experiences. He pioneered the development of methods for studying the legibility of letters. In order to study familiarity as a separate effect from legibility, we needed to use typefaces that had equal levels of legibility.

For this purpose, the first author developed four new fonts, with Spencer having the same legibility level as SpencerNeue, and PykeText the same legibility level as PykeTextNeue. The four fonts were the outcome of a previous legibility study by Beier and Larson (2010). The focus of this earlier investigation was to develop design improvements for the most often mistaken lowercase letters groups (e-c-a-s-n-u-o and i-j-l-t-f). New fonts were created as test material, each with a range of different letter variations for each of the confused characters. The letter variations were of both common and uncommon shapes. By testing these variations within each font, it was possible to locate variations that differed in commonality but had the same legibility level. Implementing these findings in two versions of the same font, one with common character shapes and one with

Familiarity conditions		Group-1 (30 participants)	Group-2 (30 participants)
Common	1) Known-common - High level of previous exposure - High level of common letter shapes	Monotype Originals Times New Roman Regular	Linotype AG Helvetica Regular
	2) New-common - Low level of previous exposure - High level of common letter shapes	Spencer Regular	PykeText Regular
New	3) New-uncommon - Low level of previous exposure - Low level of common letter shapes	PykeTextNeue Regular	SpencerNeue Regular

**Figure 3.** The fonts used to represent the three familiarity conditions.



**Figure 4.** The difference in letter shapes between the fonts Spencer and SpencerNeue (top), and Pyke and PykeNeue (bottom).

uncommon character shapes, made it possible for the present investigation to control for legibility while also varying the level of commonality. Another advantage of using newly developed fonts was that we could be sure that none of the participants had been exposed to them before the tests.

### Test material – applied fonts

The first familiarity condition (known-common) consisted of the widely-used typefaces that Frutiger demonstrates as having common letter shapes. One key member is the typeface Helvetica. As shown in Hustwit’s film *Helvetica* (2006), this is probably the most extensively used typeface today. Another frequently used typeface with common letter shapes is Times New Roman. Besides a history of being extensively 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 the faces that we, through exposure, have grown used to over

the years – for example Centaur, ITC Garamond, and Adobe Caslon.

The second familiarity condition (new-common) consisted of two newly-designed typefaces with common letter shapes that fit into the Frutiger letterform matrix. Examples of other typefaces in this category are Palatino Sans and Meta Serif, both based on standard forms, and both with famous siblings. However at the time of their release in 2005 and 2007, they were unknown to the broader public.

The third familiarity condition (new-uncommon) consisted of two typefaces with characters that appeared both novel and uncommon, with shapes differing from the Frutiger letterform matrix. Examples of these typefaces are the now unfamiliar Gothic and flourished script faces, and the various typefaces attempting to reinvent the alphabet, such as “basic alphabet” by Herbert Bayer (1967) and “New Alphabet” by Wim Crouwel.

To quantify the common letter shape of each of the fonts in this study, we computed correlations with the Frutiger letterform matrix. We looked at the amount of area inside a font that landed on black, grey, and white in





**Figure 5.** The 6 test fonts superimposed on the Frutiger letterform matrix. Reading from left to right: Helvetica, which is actually part of the Frutiger letterform matrix, not surprisingly has a high correlation of 0.73 with the letterform matrix; Times, our other known-common font, also has a very high correlation of 0.76; the new-common font Spencer has a correlation of 0.71, while the other new-common font Pyke has a correlation of 0.75. The letters a-n-s-t of new-uncommon font SpencerNeue have correlations of 0.51, while the same letters in the other new-uncommon font PykeNeue have a correlation of 0.52.

the letterform matrix. A font that overlaps heavily with the black area of the letterform matrix will have a high correlation, and a font that falls outside the letterform matrix will have a low correlation.

Figure 5 shows visually how the four common letter shape fonts all overlap the letter “a” of the letterform matrix with a high correlation of more than 0.7 (1.0 is a perfect correlation). The two uncommon letter shape fonts, on the other hand, have a letter “a” with less overlap, showing a lower correlation of about 0.5.

The test material layout was created to present the fonts with a high level of visual similarity. Since the tested fonts in general varied on x-height and width, the sizes and leading had to be adjusted internally. Pyke and PykeNeue were presented in a 10.3 point size with a leading of 16.5 point, Spencer and SpencerNeue 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., aligned to the left, with a tracking of 0, and was printed on 80 gram white A4 copy paper on a Brother HL-5220 Laser printer.

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.

**Figure 6.** Top table: Test material for Group-1. From the top: Times, Spencer, and PykeNeue.

Bottom table: Test material for Group-2. From the top: Helvetica (9.3/17), Pyke, and SpencerNeue.



## Participants

60 participants from students and staff at the Royal College of Art and Imperial College participated in this study. All test fonts were read by participants from both institutions. Participant ages ranged from 20 to 52 with an average of 28. All participants were native English speakers. All self-reported as having either normal or corrected-to-normal visual acuity, and as 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.

## Procedures

Participants completed tasks for all three familiarity conditions. The order in which the familiarity conditions were read, and the text applied in the test material, was counterbalanced (i.e. all six possible orders were used an equal number of times). Each condition consisted of a reading speed pre-test, a pre-test questionnaire, an exposure session, a reading speed post-test, and a post-test questionnaire. Participants were divided into two groups. This was done to avoid having the same participant read both **Spencer and SpencerNeue** or **both Pyke and PykeNeue**. By separating participants into two test groups, fonts of the “new” familiarity condition were guaranteed to be new to the reader.

The questionnaire was designed to elicit the reader’s opinion, which could not be quantified with a speed-of-reading test. If the focus is on performance or perception, readers’ preferences are not essential, yet when the focus is on familiarity, the subjective experience plays a central role and so makes inquiries into readers’ opinions a useful method.

In the questionnaire, 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 final 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 familiarity conditions. The reading material contained a number of short paragraphs, each with a phrase or statement towards the end of the text that stood 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 read as 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.”  
(Tinker, 1983)

For a correct response, participants had to mark the word “grocery”, after which they moved on to the next paragraph. The paragraphs were structured in such a way that participants had to 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- minute exposure session in which 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.

## Results

There was no reliable difference in reading performance between participants from the Royal College of Art or Imperial College ( $p > .05$ ). There was also no reliable difference in their questionnaire scores ( $p > .05$ ). 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 so there was no useful reason to analyze these errors.

### Reading speed of font group-1

The data shows that the average number of paragraphs read in Times (known-common condition) 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 Spencer (new-common

condition) 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 PykeNeue (new-uncommon) during the pre-test was 5.78 (SD = 2.13 paragraphs) and during the post-test 6.52 (SD = 1.85 paragraphs).

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 familiarity conditions. There was a large reliable main effect for the repeated measures 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 familiarity condition Times, Spencer, and PykeNeue,  $F(2,58)=0.70$ ,  $p > .05$ , indicating that reading speed was similar for the three fonts when averaged over the repeated measures. There was no statistically reliable interaction effect between familiarity condition and the repeated measure,  $F(2,58)=2.51$ ,  $p=.09$ .

### Reading speed for font group-2

A study of the data from font group-2 showed the average number of paragraphs read in Helvetica (known-common) during the pre-test was 5.4 (SD=1.86 paragraphs) and during the post-test 5.9 (SD= 1.74

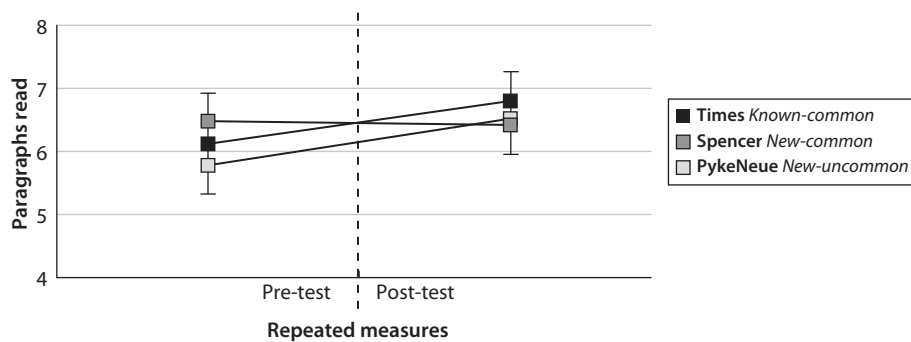


Figure 7. Average number of paragraphs read for font group-1.

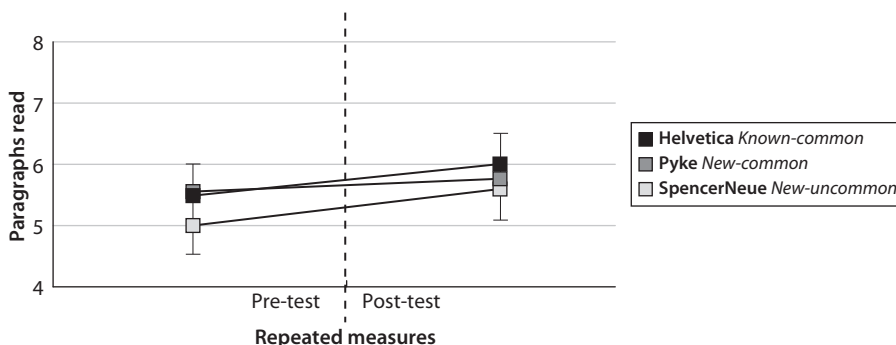


Figure 8. Average number of paragraphs read for font group-2.

paragraphs). The average number of paragraphs read in the Pyke (new-common) 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 SpencerNeue font (new-uncommon) during the pre-test was 4.9 (SD=1.69 paragraphs) and during the post-test 5.4 (SD= 1.85 paragraphs).

A 2x3 ANOVA was used to analyze the reading speed data from font group-2, showing a reliable main effect for the repeated measures 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 familiarity condition Helvetica, Pyke, and SpencerNeue,  $F(2,58)=1.80, p=.174$ , indicating that the reading speed for the three fonts was similar when averaged over the repeated measures. The interaction effect between the familiarity condition and the repeated measure was not statistically reliable,  $F(2,58)=0.92, p=.404$ .

### 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 is presented in the graphs below. The non-parametric Sheirer Ray Hare test was used to analyse the repeated measures Likert scale statements.

[Q1:] For the statement “I will enjoy reading this typeface in the future”, the known-common and new-common familiarity conditions for both font groups had medians between 1 and 2 both in the pre-test and in the post-test. The new-uncommon familiarity condition increased from a median of -2 to 0 in font group-1, and from -2 to -1 in font group-2. Font group-1 showed a reliable main effect for the familiarity condition,

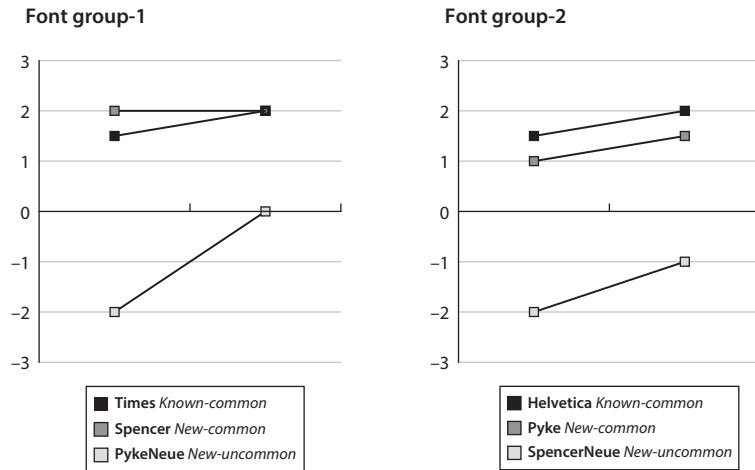
Font group	1	2
Familiarity condition	Not reliable	Not reliable
Repeated measure	Reliable	Reliable
FC x RM interaction	Not reliable	Not reliable

Figure 9. A comparison between the two font groups.

$H(2,58)=23.27, p<.0001$ . There was not a reliable measure main effect for the repeated measure,  $H(1,29)=0.75, p=.31$ , or an interaction effect between the familiarity condition and repeated measure,  $H(2,58)=2.35, p=.15$ . Font group-2 also showed a reliable main effect for the familiarity condition,  $H(2,58)=28.50, p<.0001$ . There was neither a reliable main effect for repeated measures,  $H(1,29)=1.08, p=.22$ , nor a reliable interaction effect,  $H(2,58)=1.60, p=.22$ . The results of the two font groups both demonstrated that the participants did not believe they would enjoy reading the new-uncommon typeface in the future.

[Q2:] For the statement “I was constantly focusing on the typeface”, the median scores for the known-common and new-common familiarity conditions for both test groups ranged from -1.5 to -2 both in the pre-test and

in the post-test. The new-uncommon familiarity condition decreased from a median of 2 to 0.5 in font group-1, and from 2 to 0 in font group-2. Font group-1 showed a reliable main effect for the familiarity condition,  $H(2,58)=22.17, p<.0001$ . There was neither a reliable main effect for repeated measures,  $H(1,29)=.76, p=.31$ , nor a reliable interaction effect,  $H(2,58)=2.99, p=.11$ . Font group-2 also showed a reliable main effect for the familiarity condition,  $H(2,58)=26.39, p<.0001$ . There was neither a reliable main effect for repeated measures,  $H(1,29)=.79, p=.30$ , nor a reliable interaction effect,  $H(2,58)=1.65, p=.21$ . The results of the two font groups were the same in demonstrating that participants focused more on the new-uncommon typeface than the known-common or new-common typefaces.



Font group	1	2
Familiarity condition	Reliable	Reliable
Repeated measure	Not reliable	Not reliable
FC x RM interaction	Not reliable	Not reliable

**Figure 10.** The average response to the statement, “I will enjoy reading this typeface in the future”.

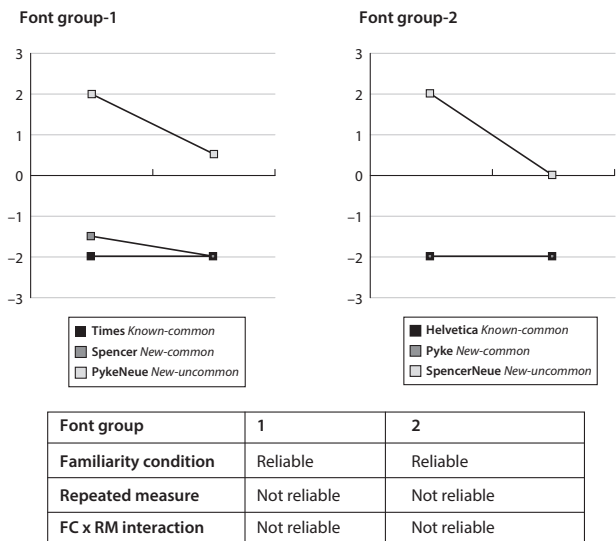


Figure 11. The average response to the statement, “I am constantly focusing on the typeface”.

[Q3:] For the statement “I still remember most of what I was reading”, the medians for most familiarity conditions ranged from 0 to 1 for both the pre-test and post-test. Only the new-uncommon condition for group-1 had a median of -1 during the pre-test. Font group-1 showed a reliable main effect for the familiarity conditions,  $H(2,58)=6.06, p=.02$ . There was neither a main effect for repeated measures,  $H(1,29)=1.65, p=.14$ , nor an interaction effect,  $H(2,58)=2.27, p=.16$ . Font group-2 showed a reliable main effect for repeated measures,  $H(1,29)=4.34, p=.02$ . There was neither a reliable main effect for familiarity condition,  $H(2,58)=1.46$ , nor a reliable interaction effect,  $H(2,58)=.90, p=.31$ . The two font groups showed different reliable effects for this questionnaire item. Participants in font group-1 reported that the new-uncommon font would be remembered less than the other two familiarity conditions, while participants in font group-2 reported that all familiarity conditions

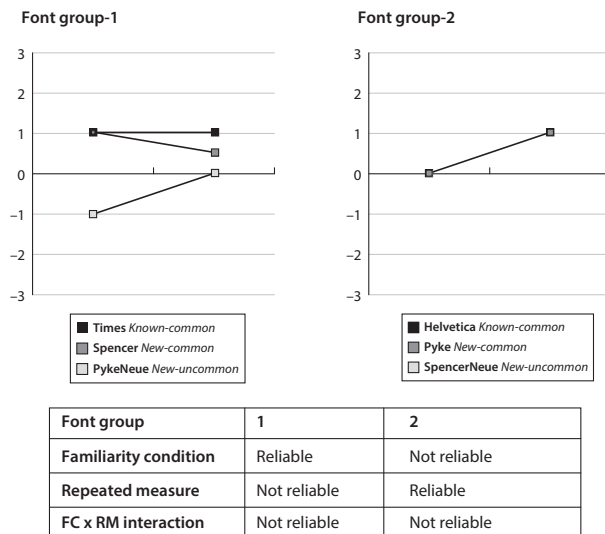


Figure 12. The average response to the statement, “I still remember most of what I was reading”.

would be better remembered after time spent reading with them.

[Q4:] For the statement “This was a comfortable reading experience”, the known-common and new-common familiarity conditions for both test groups had medians that ranged from 1 to 2 during both the pre-tests and post-tests, while the new-uncommon familiarity conditions had medians that ranged from -2 to 0. Font group-1 showed a reliable main effect for the familiarity conditions,  $H(2,58)=17.58, p<.0001$ . There was neither a reliable main effect for repeated measures,  $H(1,29)=1.25, p=.19$ , nor a reliable interaction effect,  $H(2,58)=.98, p=.31$ . Font group-2 showed reliable main effects for both the familiarity conditions,  $H(2,58)=16.02, p<.0001$ , and for repeated measures,  $H(1,29)=3.84, p=.03$ . There was not a reliable interaction effect,  $H(2,58)=1.53, p=.23$ . In both font groups the new-uncommon familiarity condition was thought to be a less comfortable reading

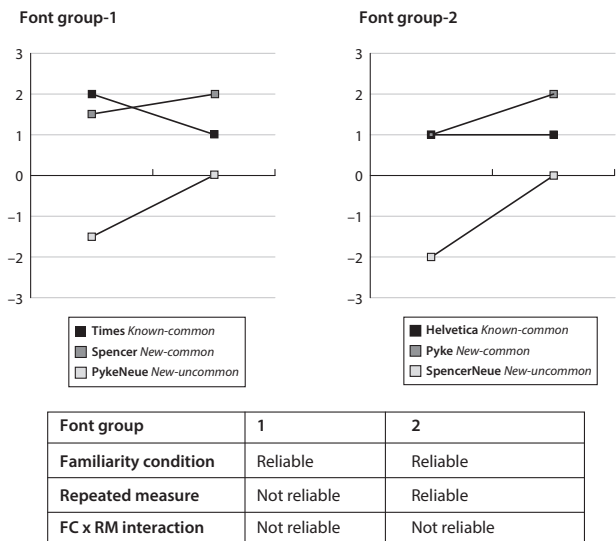


Figure 13. The average response to the statement, “This was a comfortable reading experience”.

experience, while in font group-2 participants reported an increase in comfort after the intervention.

[Q5.1:] For the statement “I have encountered this typeface before”, (only asked in the pre-test), the known-common familiarity condition had a median score of 2 in both font groups, the new-common familiarity condition had a median score of 0 in font group-1 and a median score of 2 in font group-2, and the new-uncommon familiarity condition had a median score of -3 in both font groups. A non-parametric Friedman test of ranks showed there were reliable differences in font group-1,  $Q(29)=44.54, p<.001$ , and reliable differences in font group-2,  $Q(29)=43.32, p<.001$ . Participants did not report having previously encountered the new-uncommon familiarity condition.

[Q5.2:] 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 medians for all conditions range from 0.5 to 2 for both font groups. The non-parametric Friedman test of ranks show a reliable difference in font group-1,  $Q(29)=15.37, p<.001$ , and a reliable difference in font group-2,  $Q(29)=6.68, p=.035$ . The new-uncommon familiarity conditions for both font groups were reported as highest for being easier to read now than at the beginning of the test.

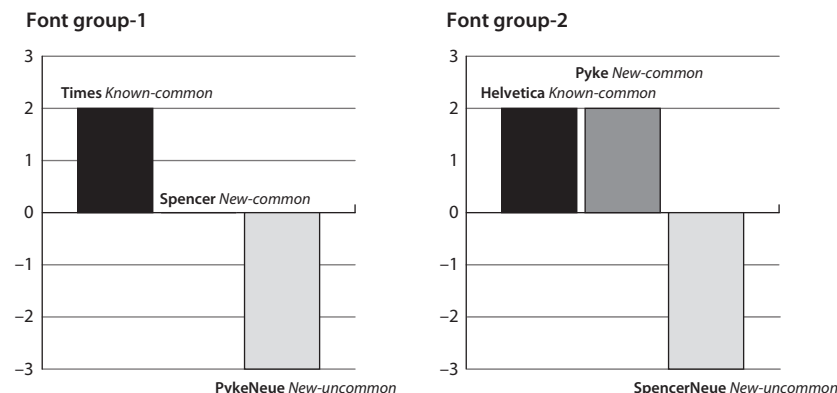
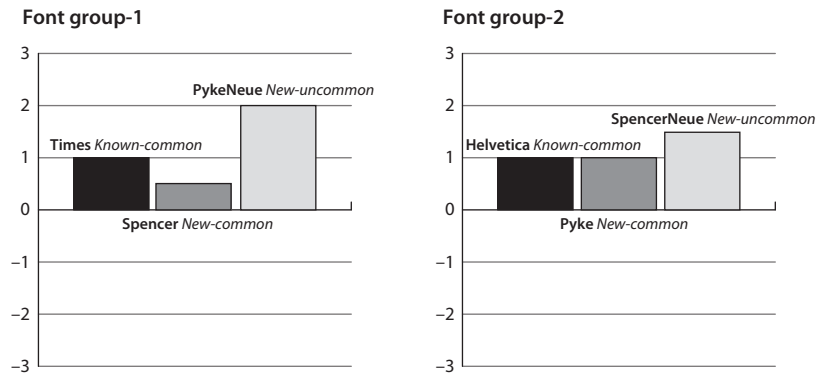


Figure 14. The average response to the statement, “I have encountered this typeface before”.



**Figure 15.** The average response to the statement, "I find the typeface easier to read now, than I did at the beginning of the test".

## Discussion

Reading speed increased from the pre-test to the post-test, indicating the presence of a learning effect. There were, in addition, a noticeably high number of reliable differences on the familiarity condition in the questionnaire data with a strong similarity between the two groups, showing that the readers' subjective impression of the fonts used in the new-uncommon condition was more critical than their impression of the two common conditions, and that they did not detect a difference in their reading of the fonts used in the known-common and new-common conditions.

### The prototype hypothesis

The hypothesis that common letter shapes are important to reading was not borne out in the reading speed performance. There was neither a main effect for familiarity conditions showing that the common fonts were different from the uncommon fonts, nor an interaction effect. However, the questionnaire data strongly supported the prototype hypothesis. For most questions there was a reliable main effect for familiarity, with the new-

uncommon condition less preferred than the common conditions.

The lack of a reading speed difference for the prototype hypothesis suggested that the prototype could be primarily a legibility effect. Most of the time, letters that fail to conform to the prototype are also less legible. But as legibility was controlled in this study, we were not seeing a statistically reliable performance effect for common versus uncommon letterforms. While there was initial dislike for the new-uncommon conditions, that dislike did decrease with exposure. We can rule out the prototype hypothesis as the cause of familiarity.

### The exposure hypothesis

The hypothesis that reading improves through exposure was demonstrated in the reading speed improvement from pre-test to post-test. There was a main effect for exposure in both font groups, where participants read at a reliably faster rate after the exposure session. But the expected interaction effect was not statistically reliable in either group.

The expected difference between the pre-test common conditions was not confirmed, showing no reli-



able interaction effect in either reading speed or readers' opinion. If previous exposure with a particular typeface is important, then the known-common fonts should be superior to the new-common fonts in the pre-test. The findings did not suggest that the level of previous exposure with a particular typeface was a factor for the reader: new typefaces may not have a negative influence on the reading process.

While there is evidence in the reading speed study to support the exposure hypothesis as the cause of familiarity, the evidence is not entirely expected, as both known-common fonts improved in performance after practice. A possible reason lies in the spacing of the fonts. Both of the known-common fonts (Helvetica and Times) have a smaller amount of space between the letters than the new fonts. The findings suggest the possibility that readers take longer to adjust to narrowly spaced fonts than to more widely spaced fonts. It is a matter that needs further research.

Another possible explanation is that exposure is not the accumulated experience with a font over a lifetime, but is something more similar to font tuning, a quick-acting effect on the visual system. Font tuning is known to take place with less than a second of exposure to a font, but perhaps it continues to develop over a slightly longer period of time.

## Conclusion

The findings of the reading speed study were unexpected. We saw a main effect for exposure confirming the belief that exposure has a strong effect on reading. Yet we saw

no main effect for the different fonts tested. The revelation that unusual letterforms do not slow down reading after a 20-minute exposure period surprisingly tells us that the level of common letterforms in typefaces is not important to reading performance. The data of the reading speed test supports the view of the proactive designers arguing in favour of improvements in letter shapes. Readers' opinions, on the other hand, support the argument of traditionalist designers by demonstrating that readers were noticeably more critical about the fonts of uncommon letter shapes compared to the fonts of common letter shapes. Based on these findings we conclude that the reason letter skeletons have changed so little over the years lies in readers' subjective opinions, and not in the way they read. In a normal reading situation, outside of the laboratory setting, readers will simply stop reading if the situation is uncomfortable, and so their opinions of the reading situation will end up overruling the prospective advantages in reading speed.

## Acknowledgment

We would like to thank Tanya Matskewich for computed letter correlations and creating Figure 1.

## Notes

1. The paper is partially based on the PhD thesis of the first author while affiliated with the Royal College of Art, London, UK.
2. Miles A. Tinker, *Tinker Speed of Reading Test* © 1947, 1955 (renewed 1983) by Miles A. Tinker. Published by the University of Minnesota Press.

## References

- Beier, S. & Larson, K. (2010). Design Improvements for Frequently Misrecognized Letters. *Information Design Journal*, 18(2), 118–137.
- Bayer, H. (1967). *Herbert Bayer: Painter, designer, architect*. New York: Reinhold Publishing Corporation.
- Dingfelder, S. F. (2006). Pretty faces: Easy on the brain? *Monitor on Psychology*, 37(9), 32.
- 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.
- Frutiger, A. (1998). *Signs and Symbols: Their Design and Meaning*. London: Ebury Press.
- Gauthier, I., Wong, A. C., Hayward, W. G., & Cheung, O. S. (2006). Font tuning associated with expertise in letter perception. *Perception*, 35(4), 541–559.
- Gill, E. (1936). *An Essay on Typography*. London: Sheed & Ward.
- Hustwit, G. (Director). (2006). *Helvetica* [film: DVD]. Plexifilm.
- Langlois, J. H., & Roggman, L.A. (1990). Attractive faces are only average. *American Psychological Society*, 1(2), 115–121.
- Licko, Z. (1990). Typeface design. *Emigre*, 15, 8–13.
- Morison, S. (1924). Towards an Ideal Type. *The Fleuron*, 2, 57–75.
- Morison, S. (1930). First Principles of Typography. *The Fleuron*, 7, 61–72.
- Sanocki, T. (1988). Font regularity constraints on the process of letter recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3), 472–480.
- Tschichold, J. (1969). Of what value is tradition in type design? In J. Dreyfus & R. Murat (Eds), *Typographic opportunities in the computer age*. Prague: Typografia, 52–55.
- Tinker, M. A. (1964). *Legibility of Print*. U.S.A.: Iowa State University Press.
- Tinker, M. A. (1983). *Tinker Speed of Reading Test*. University of Minnesota Press.
- Walker, P. (2008). Font tuning: A review and new experimental evidence. *Visual Cognition*, 16, 1022–1058.
- Zineddin, A. Z., Garvey, M., Carlson, R. A., & Pietrucha, M. T. (2003). Effects of practice on font legibility. *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting*, 1717–1720.

## About the authors

**Sofie Beier** is a type designer and assistant professor employed at the *School of Design* under *The Royal Danish Academy of Fine Arts*. She holds a PhD from the Royal College of Art in London, and is the author of the book “Reading Letters: designing for legibility”. Her research focuses on improving the reading experience by aiming at a better understanding of how different typefaces and letter shapes can influence the way we read.

Email: info@sofiebeier.dk



**Kevin Larson** received his PhD in cognitive psychology in 2000 from the University of Texas at Austin. His academic research was on word recognition and reading acquisition. His passion is understanding the impact of typography on the reading experience, and applying that understanding towards improving the on-screen reading experience. He is a member of Microsoft’s Advanced Reading Technologies team.

Email: Kevlar@microsoft.com



## Contact

Sofie Beier  
The Royal Danish Academy of Fine Arts  
School of Design  
Philip de Langes Allé 10  
1435 Copenhagen K  
Denmark

Kevin Larson  
Microsoft Advanced Reading Technologies  
1 Microsoft Way  
Redmond, WA 98052  
USA