Visible Language
58.1

An Exploratory Study Evaluating the Influence of Taller Stripe Patterns on Reading Comfort Using Ranking Tests, Reading Tests, EEG's, and Eye Tracking

Maarten Renckens
Independent Researcher

Abstract

The Latin script has a vertical stripe pattern in it, which is known to cause visual discomfort. This study started from the hypothesis that a lower stripe pattern could result in better visual comfort than a taller stripe pattern.
I evaluated this hypothesis with several letterforms and their correlating stripe patterns, tested in four independent tests: a ranking test, reading progression, measuring neurological response, and measuring eye movements. The results provide some indications that taller stripe patterns are less comfortable, but those results were mostly outside the range of common letter sizes for reading texts. Also, results for letterforms and plain stripe patterns differed. The results suggest that multiple design parameters influence reading comfort simultaneously, and that *'the number of design details per surface'* is a design parameter that could play an important role in determining reading comfort. This needs to be evaluated in further studies.

Keywords

Reading Comfort, Lower Stripe Patterns,
Number Of Details, Optimising Design Parameters

1. Research Introduction

Letters play a game of black and white, forming a repetitive stripe pattern through the repetition of black strokes. It is generally known that stripe patterns could be uncomfortable to look at, possibly arousing illusions of color, shape, and motion (Wilkins, 2012).

The effects of the stripe patterns are indicated with the term *'visual discomfort'*, the *'the subjective adverse effects encountered on viewing certain stimuli'* (e.g. O'Hare & Hibbard, 2013). In the context of research into the stripe patterns within text, I propose to use the term *'reading comfort',* which is an existing word in the Dutch language: 'leescomfort'. It is applicable to processing visual information and is commonly applied to indicate a concept as *'the best possible circumstance for reading'* (see: Vivlio, 2020; Sensotec, 2020; Denksport, 2020; Amazon, 2020; De Standaard, 2006). Because the term is not mentioned in the most important Dutch dictionaries, (Van Dale, 2015; Nederlandse Taalunie, 2001), I propose the definition: ***'The ease with which visual information can be processed.'***

While processing info, '*reading comfort'* is influenced by several factors such as the text layout (set with a clear structure, the amount of visual comfort aroused…), the content (including the already acquired knowledge/foreknowledge of the reader, being acquainted to the topic), technical aspects of writing (is it well-written), the working of the brain (learning disabilities, brain injuries…), the environment (disturbing noises)… (this list is extended from Duin, 2013). In my definition of reading comfort, an easy processing of information correlates with a high comfort.

There is not much information about the relation between the stripe pattern in letters and their reading comfort. Stripe patterns are present in two directions: horizontally through the lines of text (e.g. Wilkins, 2012) and vertically through the vertical strokes in the letter shapes (e.g. Wilkins et al., 2007). In text, stripe patterns can hinder readers sensitive to visual stress such as readers suffering from migraine (Wilkins, 2007), and can cause epileptic seizures (Harding et al., 2017; Wilkins, 2012). Stripe patterns in common visual input (TV broadcasts, games) are regulated to avoid such negative effects (British Independent Television Commission, 2001). But regulation does not exist for letters, even though stripe patterns are an inherent part of Latin letterforms.

Regarding the stripe pattern, earlier research on letterforms focused entirely on the horizontal and unequal distribution of the stripe pattern, but none focused on taller letters. This is a pity. Several letters from the Latin alphabet are bound to their strong one-directional pattern, but the height of the stripe pattern just happens to be an independent design parameter which a designer can change easily. Other parameters of the stripe patterns that can be changed are the space between the vertical letter strokes and their thickness. Even more, each individual letter, each typeface, each font (a variation on one typeface) and each individual letter already possesses a stripe pattern of its own. For instance, a bold font has a smaller distance between the vertical strokes and thicker vertical strokes than the light font within the same typeface. Designers and researchers did experiments on the horizontal distance between vertical stripes. It came forward that words with a rigid stripe pattern take a longer reading duration than words with a less rigid stripe pattern such as present in rounded letters, such as the word "mum" versus the word "dad" (Wilkins et al., 2012; Jainta, Jaschinski & Wilkins, 2010). A less evenly spaced stripe pattern improves legibility for children with a visual impairment (Bessemans, 2012) and increases the speed of word recognition (Wilkins et al., 2007). However, there are also parameters which a type designer cannot change. For example, in an experiment (not involving letters) from Pennachio and Wilkins (2015), viewers favor visual input wherein the stripe pattern is set up in several (overlapping) orientations. But applying this to letters would limit the letters' legibility and would look very sloppy.

Studying the stripe pattern in letters has the purpose of increasing the 'reading comfort'. In earlier studies, the stripe pattern was defined as 'the component strokes' or as 'the sequence of strokes' (Bessemans, 2012; Wilkins et al., 2007). This does not clarify the exact position of the stripe pattern in letters [also named the rhythm]; and specifically, not its height. A clear definition was needed to continue. The following definition was considered viable: 'The rhythm in type is the sequence of the longest continuous black masses within the letters, in any direction' (Renckens, 2020). This definition consequently positions the stripe pattern, starting from above the bottom serifs if present, up to the letters' counter forms (the white). Based on this definition, the stripe pattern has several different occurrences in different letter shapes [Figure 1]. On those differences on which this study focusses.

2. Hypothesis and Research Setup

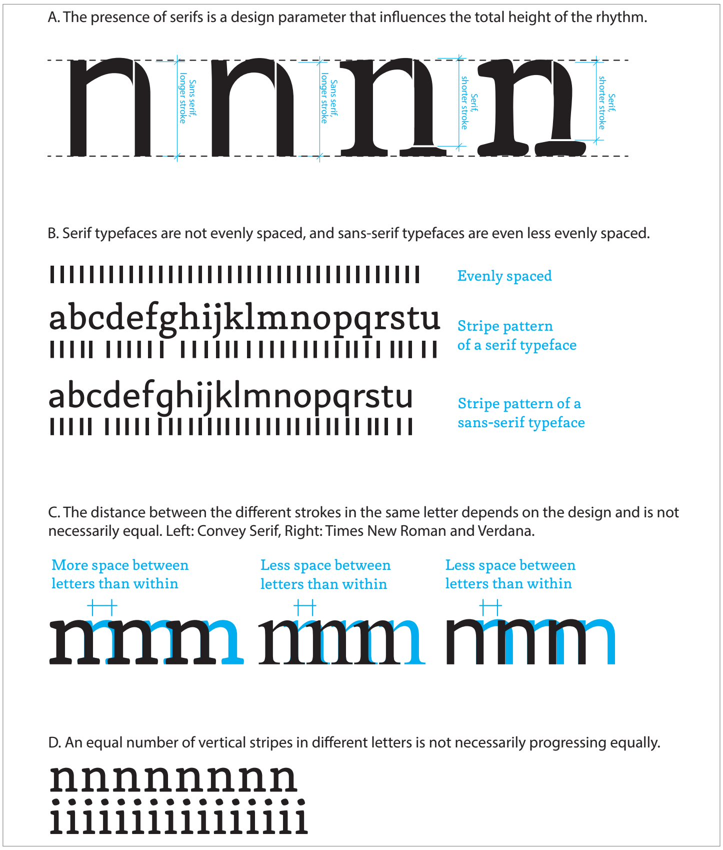
To improve the reading comfort of letters, more knowledge about the effect of the stripe pattern in letters on the reading process is needed. The current study primary focuses on how reading comfort is influenced by taller stripe patterns. A secondary focus is on how more condensed and less evenly spaced stripe patterns influence reading comfort. This study starts from the following paradigms:

1 It is known that stressful visual information results in a changed neurological response, including spiked waveforms (Wilkins, 2012). This causes visual discomfort when visual stimuli provoke an extremely strong neural response (O'Hare & Hibbard, 2011). Taller and more condensed stripe patterns are more prominent, which could result in more neurological response.

2 A taller stripe pattern has more vertical distance from top to bottom. Therefore, it could cause more eye closures and require more time before a line of text is scanned with the eyes than a smaller stripe pattern.

FIGURE 1:

Stripe patterns differ between several typefaces and letterforms.



i *The height of the stripe pattern is taller in sans-serif typefaces [Figure 1a]. This is because serifs shorten the vertical stripes (Renckens, 2020);*

ii *The stripe pattern is more condensed in some letters of sans-serif typefaces than within the same letters of serif typefaces [Figure 1b]. This is because serifs require additional space between the letters, in order not to touch each other.*

iii *The stripe pattern in a serif typeface is more evenly spaced and the stripe pattern in a sans-serif typeface is less evenly spaced, but neither are completely evenly spaced [Figure 1b]. This is again because serifs require additional space between the letters, in order not to touch each other.*

iv *For letters with a vertical stroke, the space between consecutive letters is in most cases smaller than the space within letter shapes [Figure 1c].*

3 Crowding (a limited recognition of objects further away from the eye's fovea due to neighboring objects close by) relates to closely spaced contours (Jainta, Jaschinski & Wilkins, 2010). Therefore, it could be that horizontally more condensed stripe patterns in sans-serif typefaces require more time before a line of text is scanned with the eyes than the less condensed stripe pattern in serif fonts.

4 The effects from (1), (2) and (3) could cause a less efficient reading progression, resulting in a slower reading pace.

Based upon the paradigms, the hypothesis is stated that "taller, less evenly spaced, and more condensed stripe patterns could be less effective for the reading comfort." If the hypotheses could be confirmed, this research will not only provide guidelines for type design with more reading comfort but could also contribute to answer the frequently asked question 'which is better: a serif or sans-serif typeface?', for as far as a simple answer can ever be provided on this question.

In this study, the hypothesis is tested with four tests, each focusing on a different methodology and with different research questions (discussed later). The basis for all tests is the typeface "Convey," designed for reading text by the author and adjusted to contain a total of 19 different fonts [Figure 2]. Some of these fonts contain letters, while others represent the stripe pattern within the letters.

The main difference between the fonts is the height. The number for each font's name refers to certain measurements used in type design (UPM, Units per Em). All fonts containing letters are designed to respect internal and external research validity. This means that on the one hand, only one parameter changes (the height of the stripe pattern), ensuring that valid comparisons can be made, but on the other hand, each font is individually re-designed instead of just being vertically scaled. This avoids disproportion and ensures that the fonts are suited for real-life reading materials. Some special fonts were added: a horizontally evenly spaced stripe pattern {R700reg} and two '700-extra' fonts, fonts with a strikethrough {Rsans700_mid and Rserif700_bot}. Those '700-extra's' do not refer to letterforms but serve to test the visibility of stripe patterns when they are shortened by an interruption.

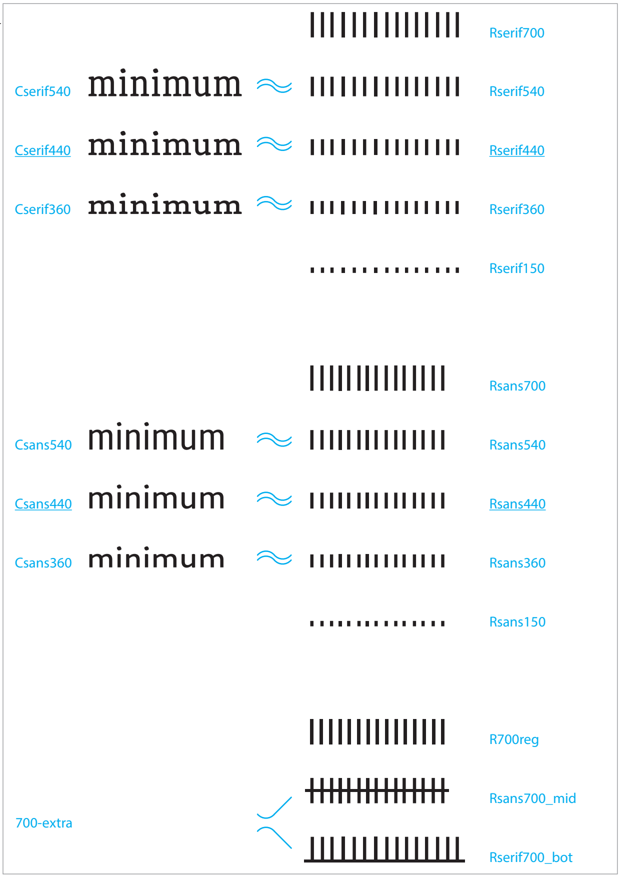
Because not all letters of the Latin alphabet possess the same stripe pattern, the test materials of this research mostly focus on the letters {nmiu}, which contain straight strokes. For some tests, additional letters are required. Those will be described in the chapter about the test itself.

In all test materials, attention is paid to how the multiple text lines were set. It is known that lines of text also form a horizontal stripe pattern, possibly causing visual discomfort. It is known that readers prefer more line spacing in the layout (Wilkins & Nimmo-Smith, 1987). In the testing materials for this research, text is set with a slightly larger leading (130%) than the often-used default (120%). That extra spacing ensured that the fonts with a larger x-height did not come too close to each other. This line height of 130% was kept throughout this whole research.

For this research, most participants are random students at the university where the lab is situated. Some participants who couldn't make their way to the lab participated at another place in the tests 1 and 2, because those tests didn't require the eye tracking camera or the EEG-cap. A fee was paid for participating, independent of the answers on the tests. The only requirement for participants is that they needed to have normal or corrected to normal eye vision. Regarding the language, non-Dutch speakers are excluded from the reading aloud test; but could participate in all other tests. Learning disabilities are not a base for exclusion but are registered and added to the database for analysis in possible follow-up studies.

FIGURE 2:

All 19 fonts of the typeface 'Convey' (a 'font' is a variation on a typeface). The number indicates the height, with the number 440 correlating with a common x-height.



3. Test 1: Ranking According To Visual Comfort

3.1. Test 1: research questions

1 Is a taller stripe pattern in text perceived as less favorable to look at, compared to lower stripe patterns?

2 Is a more condensed and more unevenly spaced stripe pattern in text perceived as less favorable to look at, compared to wider stripe patterns?

3.2. Test 1: research materials

For the first test, the fonts from Figure 2 are set in the word minimum, in a 5x5 grid [Figure 3]. All 19 fonts are printed on separate papers in high definition. The 13 fonts with stripes form the first series and the 6 fonts with letters form a second series.

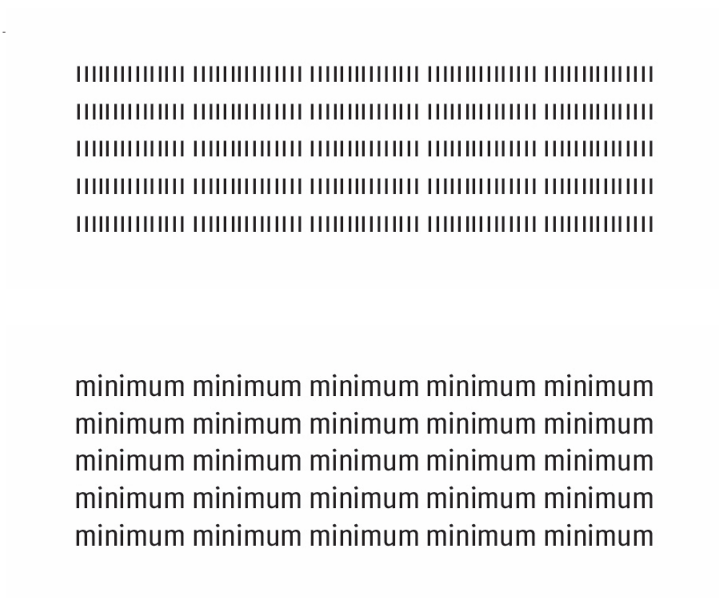
3.3. Test 1: methodology

A total of 51 persons participated in this test. For this test, participants were first asked to rank the 13 sheets of paper containing stripe patterns, and thereafter to rank the 6 sheets with letters. There was no time limit on the test.

All sheets were put on a table by the researcher in a random way, not in a straight line, so participants didn't receive a clue about the expectations of the research.

FIGURE 3:

Test materials for test 1 consisted of 13 papers with different stripe patterns and 6 papers with different letters.



3.4. Test 1: data collection

Participants were asked to rank the stripe patterns from most comfortable to look at to the least comfortable to look at. If a participant found this terminology confusing, the questions were rephrased as "Which ones could you easily look at for a long time" and/or "Which ones hurt your eyes." After the participant ranked the stripe patterns, a conversation was started about the order of the stripes. Some participants told me that they tried to read the stripe pattern or made certain associations with it, which was obviously not the intention: some recognized the word 'minimum' in it, others preferred the strikethrough pattern because it looked like the drawings on a prison wall. In those cases, when a participant clearly diverged from the intention, those participants were asked to do the test again concentrating on the feeling the stripe pattern provided them. In case a participant said that some stripe patterns were 'clearer' than others, they were also asked what they meant. In most of those cases, participants then replied that they could read the stripe patterns better. Those participants were also asked to do the test again and to concentrate on their feelings.

After the stripe patterns were ranked, participants were asked to rank the 6 sheets with the letters from the most comfortable to read to least comfortable to read. The assignment was clarified with the explanation "try to imagine you need to read a very long text with each typeface. Which one would you prefer in that case?"

3.5. Test 1: results

The following tables all illustrate the results from more comfortable' (left) to 'less comfortable' (right). Tables 1 and 3 list the fonts individually, tables 2 and 4 groups them per similar height plus the '700-extra's' separated. The expected place of each font in the ranking is indicated by a border. Lighter background colors indicate that the font is often ranked in this position. The lower fonts were ranked more often as 'more comfortable' (left), the taller fonts were ranked more often as 'less comfortable' (right). Notice that the '700-extra' fonts were ranked as the least comfortable to look at.

3.6. Test 1: analysis

A Friedman test was used for the analysis of the difference in the ranking of the fonts by the participants with α=0.05. A rejection of the null hypothesis means that the ranking of at least one font is different.

A post-hoc paired comparison of the ranking of the fonts via the Eisinga exact test with α=0.05 corrected with the

TABLE 1:

For the individual stripe patterns, participants ranked the smaller stripe pattern as more comfortable (left) and taller stripe patterns as less comfortable (right); but the results are not always very distinct and the two '700-extra' fonts are ranked the least comfortable, against the expectations. Each stripe pattern had 51 occurrences.

	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6	Rank 7	Rank 8	Rank 9	Rank 10	Rank 11	Rank 12	Rank 13
Rserif150	55%	14%	8%	2%	4%	4%	2%	2%	4%	2%	2%	2%	0%
Rsans150	14%	55%	0%	6%	6%	2%	2%	6%	2%	2%	6%	0%	0%
Rserif360	6%	4%	35%	16%	8%	4%	4%	10%	0%	0%	2%	10%	2%
Rsans360	2%	4%	8%	25%	8%	10%	10%	4%	8%	6%	0%	4%	12%
Rserif440	0%	8%	16%	16%	14%	20%	8%	6%	6%	4%	4%	0%	0%
Rsans440	0%	2%	0%	10%	18%	24%	10%	6%	14%	6%	10%	2%	0%
Rserif540	12%	4%	12%	4%	16%	8%	24%	12%	6%	2%	2%	0%	0%
Rsans540	0%	2%	2%	6%	6%	8%	18%	33%	12%	4%	10%	0%	0%
Rsans700_mid	0%	2%	2%	4%	0%	2%	0%	2%	2%	10%	6%	10%	61%
Rserif700_bot	4%	0%	10%	4%	2%	0%	4%	6%	6%	10%	2%	43%	10%
Rserif700	2%	6%	2%	6%	10%	4%	10%	6%	29%	10%	10%	6%	0%
Rsans700	0%	0%	4%	2%	6%	10%	2%	8%	4%	27%	22%	16%	0%
R700reg	6%	0%	2%	0%	4%	6%	8%	0%	8%	18%	25%	8%	16%

TABLE 2:

Also, when grouped per similar height, participants ranked the smaller stripe pattern as more comfortable (left) and taller stripe patterns as less comfortable (right); but the two '700-extra' fonts are ranked the least comfortable, against the expectations. Each stripe pattern had 102 occurrences, except for the 700's which had 153 occurrences in the group.

	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
150-group	69%	8%	8%	6%	5%	5%
360-group	8%	42%	15%	14%	7%	15%
440-group	5%	21%	37%	15%	15%	8%
540-group	9%	12%	19%	43%	12%	6%
700-extra	3%	10%	2%	6%	14%	66%
700-group	5%	5%	13%	11%	32%	34%

TABLE 3:

Participants ranked the middle size letters {440} as most comfortable (left) and the {360} and {540} sizes as less comfortable (right). Each font with letterforms pattern had 51 occurrences.

	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
Cserif360	2%	14%	18%	14%	27%	25%
Csans360	12%	8%	16%	16%	25%	24%
Cserif440	35%	24%	18%	20%	4%	0%
Csans440	33%	27%	16%	20%	4%	0%
Cserif540	10%	12%	20%	16%	27%	16%
Csans540	8%	16%	14%	16%	12%	35%

TABLE 4:

Also, when grouped, participants ranked the middle size letters {440} as most comfortable (left) and the {360} and {540} sizes as less comfortable (right). Each stripe pattern had 102 occurrences.

	Rank 1	Rank 2	Rank 3
360-group	18%	31%	51%
440-group	60%	36%	4%
540-group	23%	32%	45%

Bonferroni method compares if the ranking of each pair of the fonts is different and allows the identification of the differences in the fonts in combination with the summed ranks. A rejection of the null hypothesis means that the ranking of at least one font is different.

The Friedman test resulted in p-values of <0.0001, meaning that there is a difference between the ranking of the fonts. The mid-rank is used to rank the aggregated groups.

Analyzing the 13 fonts with stripe patterns

Rserif150 and Rsans150 are ranked statistically significant more often as 'most comfortable'. Ranking the fonts by the summed rank over the subjects shows the following order (a lower summed rank means ranked as 'more comfortable'): Rserif150 (146) > Rsans150 (179.0) > Rserif360 (263.0) > Rserif440 (280.0) > Rserif540 (280.0) > Rsans360 (338.0) > Rsans440 (358.0) > Rsans540 (384.0) > Rserif700 (394.0) > Rsans700 (469.0) > Regular (483.0) > Rserif700_bot (485.0) > Rsans700_mid (582.0). There is no difference in ranking between those two 150 fonts. The eleven other fonts are ranked statistically significant 'less comfortable' than the 150 fonts, but there is no difference between these eleven fonts.

But when grouping fonts with similar height together, the 150-group is again significantly more often ranked as 'most comfortable', but the 700-extra's group is now significantly more often ranked as 'least comfortable'. The summed rank here is R150 (162.5) > R360 (300.5) > R440 (319.0) > R540 (332.0) > R700 (448.7) > 700-extra's (533.5), in which a lower number means ranked as 'more comfortable'. There is no statistically significant difference between the 360, 440, 540 and 700-groups, which are positioned in the middle.

Analyzing the 6 fonts with letters

Csans440 and Cserif440 are statistically significant more often ranked as 'more comfortable' than the other fonts. This results in the summed rank is Csans440 (119.0) = Cserif440 (119.0) > Cserif540 (197.0) > Csans360 (207.0) > Csans540 (211.0) > Cserif360 (218.0), in which a lower number means ranked as 'more comfortable'. There is no difference in ranking between those two 440 fonts. There is no statistically significant difference between the other four fonts.

When grouping fonts with similar heights together, the results are the same: the 440-group is statistically significantly more often ranked as 'more comfortable' than the other groups. The summed rank is C440 (119.0) > C540 (204.0) > C360 (212.5), in which a lower number means ranked as 'more comfortable'. There is no statistically significant difference between the other two groups.

Participants' comments

Participants' comments were in line with the results. For the stripe patterns, participants indicated that the lower stripe patterns caused some order, some rest, while the taller stripe patterns and the stripe patterns with strike throughs were restless. For the letterforms, participants pointed out that the middle sizes were the ones they were used to, and that the other sizes were strange, and sometimes too crowded.

Only one participant differentiated and expressed 'the more evenly spaced fonts' (serif) versus 'the less evenly spaced fonts' (sans-serif). She said that she was an interior architect and the less evenly spaced stripe patterns were 'not nice to look at.'

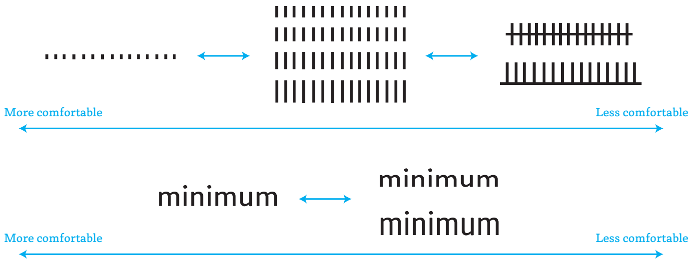


FIGURE 4:

Graphic visualization of the results. The smallest stripe patterns are ranked significantly more often as comfortable than all other stripe patterns. When the stimuli are grouped together per height, also the 700-extra's-group is significantly more often ranked as less comfortable. For the letters, the middle size was ranked more comfortable than the taller or smaller letters.

3.7. Test 1: conclusions

There are significant differences in the ranking of stripe patterns and letters from the least favorable to the most favorable.

Participants were sensitive to smaller stripe patterns, even if they are only a couple of millimeters high. The smallest stripe patterns ranked most comfortable. However, the visual differences between all the stripe patterns were probably too small, participants had some trouble distinguishing the differences. The (non-significant) results show a tendency that how taller the stripe patterns become, the less comfortable they are ranked. Doing this test again with lesser and more distinct stimuli will probably result in clearer results.

What strikes most is that the 700-extra's, the stripe patterns with a strikethrough, are ranked least comfortable, even less than the 700-straight fonts. While I hypothesized that a strikethrough would break the stripe pattern, making it more comfortable, it seems that the stripe patterns with a strikethrough are the least preferred. It seems that having more details in sight is less preferred than having a tall stripe pattern in sight. A future study can dive deeper into this to clarify those results.

This test clearly proves that the findings for the stripe patterns are not applicable on letters. For the letters, the middle size was ranked as the most comfortable. The results could not clarify why this occurred. This again is probably because there are more details to process. Future studies will clarify.

4. Test 2: reading progression in 45 seconds

4.1. Test 2: research questions

1 Does a taller stripe pattern in text cause a slower reading progression compared to lower stripe patterns?

2 Does a more condensed and more unevenly spaced stripe pattern in text cause a slower reading progression compared to wider stripe patterns?

4.2. Test 2: research materials

This test applied the letterforms {Cserif360, Cserif440, Cserif540, Csans360, Csans440, Csans540}. For this test, each font was applied to a list of random words and pseudo-words, half a page long [Figure 5]. Each word contained only one syllable. The fact that all words were of equal length enabled counting the number of words read within a specific time span.

The test was designed in such a way that learning effects were avoided. First, the test was designed four times, each time with a different random order of the fonts. Second, the list of words on a page was long, avoiding that participants would memorize words and thus unintendedly would increase the reading speed.

FIGURE 5:

Test materials for test 2 is a reading text consisting of random words from one syllable

nep diep gig dool blij pel peel gum eb hiel mop buil gun gong
lui lijm lig moe end dood glooi hulp dub gin loop glom bel moei
gil elp zes hooi dop bed je neep huld bid held pijn een leb in
blond mog sol om nul boog bloei tien fa led pep doog het duim
hield goed pol ing ei jong eb blup muil god lijd beul puin del
bon leun dun geel oud leid rel peld acht gilt hond pel beun iel
doom boom pep jij lip bug mob luid bui hup don dood len nen
hoen ju nog pi hen hop kon plooi gem pulp pijp heil joop deel

4.3. Test 2: methodology

There were 26 participants who participated in this test. Non-native Dutch speakers were left out because language was a parameter that could influence the results.

The researcher shuffled the sheets before the test began, so he did not see the order in which the participant would read the texts. Participants read aloud each of the six pages for 45 seconds. The reading progression was registered by the researcher, who counted the number of words a participant read and the number of reading errors on each of the six pages. The test took more or less five minutes per participant.

4.4. Test 2: data collection

The reading test measures the reading progression of readers by counting the number of words read aloud and the number of errors while reading within a timespan of 45 seconds.

4.5. Test 2: results

Each font was read once by each participant, so each font was read a total of 26 times. Table 5 lists the 6 fonts with letters individually and table 6 lists the 6 fonts with letters grouped per similar height.

TABLE 5: *The results of the reading aloud test, with all fonts individually. Lighter background colors indicate the 'best' reading achievements.*

	# words r	av. words read		# errors	av. Errors
Cserif360	2058	79.154		34	1.65%
Csans360	2085	80.192		29	1.39%
Cserif440	2097	80.654		30	1.43%
Csans440	2160	83.077		35	1.62%
Cserif540	2006	77.154		38	1.89%
Csans540	2047	78.731		31	1.51%

TABLE 6: *The results of the reading aloud test, with the fonts grouped per similar height. Lighter background colors indicate the 'best' reading achievements.*

	# words r	av. words read		# errors	av. errors
360	4143	79.673		63	0.01521
440	4257	81.865		65	0.01527
540	4053	77.942		69	0.01702

4.6. Test 2: analysis

A generalized linear mixed model with Tukey's correction for (pairwise) multiple testing (p=0.05) was used to test if the {number of read words} or {the number of reading errors} differs between the fonts, considering the correlation of the reading test for the same participant. A rejection of the null hypothesis means that there is a difference between the number of read words or the number of reading errors between the fonts. The pairwise comparison of the number of read words per font provides two statistically significant differences [Figure 6]:

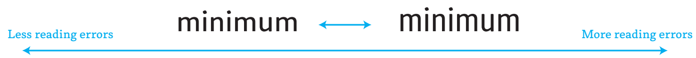
- In Cserif540 are less words read correctly than in Csans440 (p=0.0077).
- When grouping letters with similar height, group C540 has less words read correctly than in group C440 (p=0.0036).

4.7. Test 2: conclusions

In the taller fonts, statistically significant more reading errors are made. This was expected and suggests that there is an effect of taller stripe patterns on reading.

FIGURE 6:

In all cases that provided significant results, the C540 fonts were read with more reading errors than in the C440 group. The difference in reading speed was not significant.



The (non-significant) results provide additional info which is worth looking into. The 440-group allows readers to read more words aloud within a specific time span. And the fonts with which the most words are read aloud, are not read with the least errors. However, those results are not statistically significant in the current dataset. On this, more research is needed.

5. Test 3: EEG recordings of neurological response while looking at taller/more condensed letters

5.1. Test 3: research questions

1 Does a taller stripe pattern in text induce more neurological response compared to lower stripe patterns?

2 Does a more condensed and more unevenly spaced stripe pattern in text induce more neurological response compared to wider stripe patterns?

5.2. Test 3: research materials

This test evaluates if the neurological response changes for different stimuli. The screen was a 24-inch VIEWPixx /EEG with a resolution of 1,920 x 1,080 pixels.

Subtest 1a

For the first EEG subtest, visual stimuli were made by setting the word 'minimum' in all 19 typefaces from Figure 2. The visual stimuli were flashed on the screen for 300ms [Figure 7]. The short duration was intended to avoid eye movements, and thus to avoid additional neurological response. Each image was preceded by a blank screen that was visible for 2,000ms with a dot in the middle. The dots allow the participant to keep the eye at the center of the screen between different stimuli.

FIGURE 7:

Zoomed-in example materials for the first subtest. The word 'minimum' is flashed for 300ms in the 19 different fonts, containing both letters and stripe patterns. For the actual test, the colors were reversed to white text on a black background.

Subtest 1b

This test evaluates if the neurological response changes when stimuli are shown for a longer time period. For the second EEG subtest, visual stimuli were composed of five lines of text set in all 19 fonts from Figure 2 [Figure 8]. Only the word 'minimum' was used in these lines of text, to avoid possible different associations with different words, and thus different brain activities. The stimuli followed upon each other with a duration of 2,000ms, and in this subtest there were no dots between the different stimuli.

Subtest 2

This test evaluates if the neurological response changes when text intended to read silently is set in different fonts. In this subtest, again the typeface Convey was used, but only the six fonts containing letters. Additionally, four other fonts were added, containing longer ascenders/descenders (parts below the baseline and above the x-height: bdpqj) to reach a more visible stripe pattern. Those fonts are indicated by the term 'min' or 'plus' in their names in Figure 9.

To present those fonts to the participants, sentences with random words were created. The same words were used every time, ensuring the same difficulty level on every page (and thus for every font), but the order of the words was mixed every time in order to avoid learning effects. When a page was read, the participant pressed a key to go to the next font.

FIGURE 8:

Zoomed-in example materials for subtest 1b containing a grid of 5 words long and five lines. The top image shows the assumed visually most stressful stripe pattern and the bottom image shows the assumed visually most stressful letters that were presented to the participants for 2,000ms. For the actual test, the colors were reversed to white text on a black background.

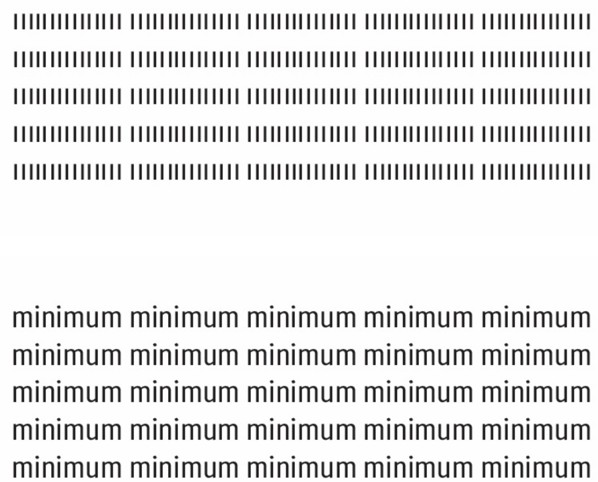
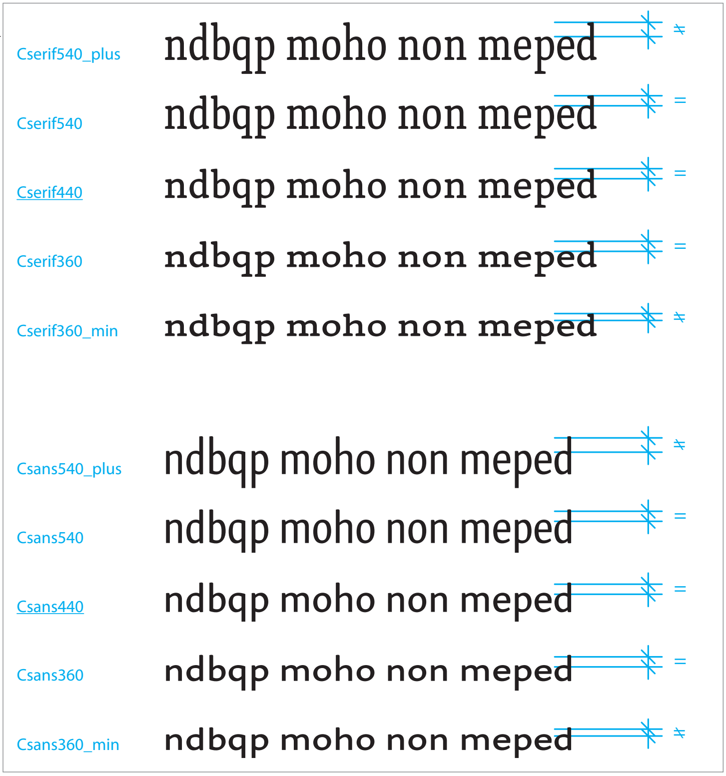


FIGURE 9:

Subtest 2: the testing fonts, with differences in the ascender and descender lengths on the plus and min fonts, to further increase/ decrease the length of the stripe pattern in a way that is commonly applied in type design.



5.3. Test 3: methodology

A total of 42 persons participated. The test lasted about an hour per participant, including half an hour of attaching the EEG cap to the scalp wet gel and the trials for testing the connection. During the set-up, participants were put at ease with some small talk about daily-life subjects.

For subtests 1a and 1b, the participant was asked to look carefully at the screen. To keep participants' attention, they were told to look/read with attention, because they would receive questions about what they were to see. For test 2, participants were told to look to the screen and to read with attention and at a normal speed, and that they should vocalize each word in their head. To keep their attention, they were told that a question would follow at the end.

The three subtests were shown to the participants in a random order, so not every participant started with test 1A. At the end of each subtest, participants were asked if 'they noticed something, followed by open questions inviting them to describe their perception and possible hindrances encountered during the test. To keep the participants' attention, the visual stimuli 'Cserif440' and 'Csans440' were shown twice, but with the dot of one letter 'i' moved to the right. Inserting this small eye-catcher was done so every participant would have an easy starting point in the conversation later on, when they were asked to describe the things that they noticed within the visual input. Only if a participant started commenting about the fonts/letters, more focused questions were asked; for example, to describe the noticed design features. If participants didn't start talking about any design features, they were not pointed to any.

In the physical setup, the screen always had a distance of 60 cm from the participant's forehead. This distance was chosen so the whole screen was still in view, and not only partially (and because the reading distance from a screen differs from the reading distance from paper which is more or less 40cm: Legge & Bigelow, 2011; Wilkins et al., 1984). Once participants confirmed that their posture was optimal for the reading test, participants were asked to place the chin on a chin holder, not to re-position the chair and not to change their posture. Before a participant started the tests, a screen containing a welcome screen was shown and the question was asked if the participant could read the words easily.

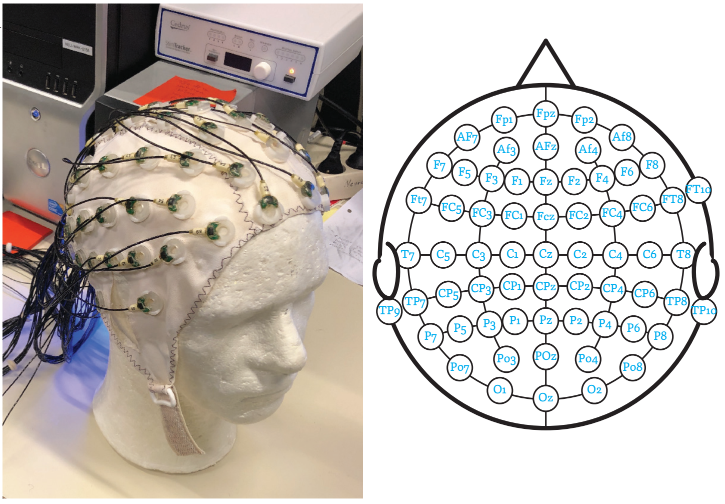
All tests (1a, 1b, and 2) were shown four times to each participant. To avoid learning effects, the computer randomized all the visual input each time it was shown. This also means that every participant looked at all stimuli four times.

5.4. Test 3: data collection

The neurological response was registered with 64 electrodes on a 1,000Hz frequency [Figure 10]. EEG signals were recorded with a 64-channel synapse system and digitized with a sampling rate of 1,000 Hz. All preprocessing and analysis were done with MNE-Python, an

FIGURE 10:

The EEG cap with 64 individual electrodes and their positions.



open-source Python package for exploring and analyzing neurophysiological data (Gramfort et al. 2013). Noisy channels were identified with visual inspection and interpolated. The data was then filtered with a zero-phase band-pass filter between 1Hz and 60Hz. An additional notch filter was applied at 50 Hz in order to remove electrical line noise. Independent component analysis (ICA) was used to identify ocular and movement artifacts. An automated routine for detection of ocular components in the ICA was performed based on the EOG signal. The components identified by this routine were subjected to a post-hoc manual check for additional non-brain signals. The resulting ICA solution was applied to the filtered signal.

The continuous EEG signal was then divided into epochs using stimulus onset triggers. For tests 1a and 1b, the epochs were 2 seconds long while for test 2 the epochs were 15 seconds long. Participants could read as long as they wished, but I had to take equal lengths for the analysis. The shortest duration that any participant did over reading the stimuli in test 2 was used, namely 15 seconds. In each case, a baseline correction was applied with a baseline period of 100ms. Invalid epochs were rejected based on signal amplitude. Overall, 2,678 epochs were included in the analysis for subtest 1a, 2,750 for test 1b and 333 for test 2. Fourier spectral analysis was performed on each epoch using DPSS tapers. Spectral power was averaged for the whole scalp, and then compared across each frequency bin. The frequency bin width was 0.48 Hz, as determined by the overall signal bandwidth and sampling rate. For pairwise comparisons between conditions, t-tests with Welch's correction for unequal variances were used.

5.5. Test 3: results

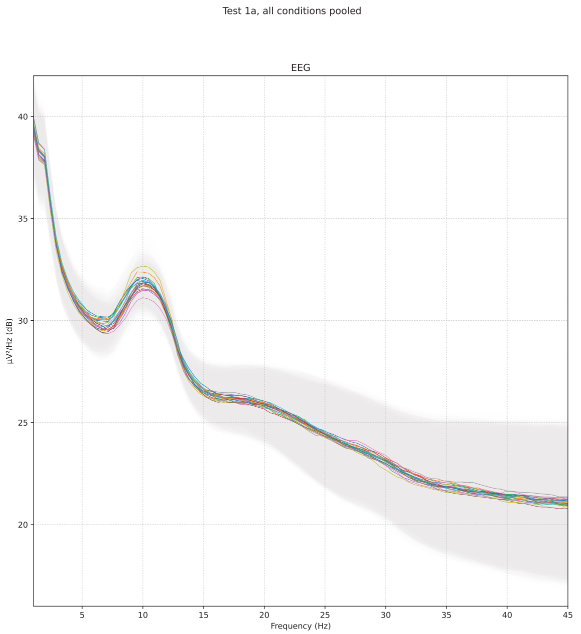
The data of 2 participants was lost due to technical errors and were left out. 40 participants remained. Each participant looked at the visual stimuli 4 times, which means that there are 164 stimuli in this dataset. Figure 11 provides an idea about how to imagine the dataset.

5.6. Test 3: analyses

For this study, I am very much aware of the multiple comparisons problem which requests to inflate p-values when many comparisons are being made. The multiple comparisons correction was not applied because the current study is exploratory, meaning that a lot of different stimuli are included to detect what directions to go with future research. Later studies need to examine the results in more detail and will probably have less comparisons to be made. Therefore, it is not useful to apply stricter rules now than would be done in later studies.

FIGURE 11:

The EEG cap with 64 individual electrodes and their positions.



Subtest 1a: Statistical analysis

For subtest 1a, the test in which all participants looked to the visual stimuli for 300ms, no statistically significant differences ($p<0.05$) were found between the visual stimuli.

Subtest 1b: Statistical analysis

For subtest 1b, the test in which all participants looked to the visual stimuli for 1,000ms, statistically significant differences were found between some frequencies in the neurological response, between specific visual stimuli [tables 7]. The tables are included in the article to show a truthful view of where statistically significant results are present, but also to show where they aren't.

When analyzing all stripe patterns individually [table 7a], taller stripe patterns were associated with more neurological response in all statistically significant results. The only exception is the Rsans360 vs Rsans440 comparison, where Rsans360 produced higher responses. When analyzing the stripe patterns per similar height [table 7b], R150 was associated with smaller responses than '700-extra'. When analyzing the letterforms, the 360 fonts yielded smaller responses in all comparisons where a statistically significant effect occurred [table 7c and d].

TABLE 7(a–d):

The tables indicate which comparison of visual stimuli resulted in a statistically significant difference. The numbers specify the frequencies at which those differences occurred. So, if a cell contains [5.24 5.71] for example, it means that differences were found for frequency bins at 5.24Hz and 5.71Hz. Each frequency bin has a width of 0.48 Hz. Groups are as following: R150 = {Rserif150, Rsans150}, R360 = {Rserif360, Rsans360}, R440 = {Rserif440, Rsans440}, R540 = {Rserif540, Rsans540}, R700 = {Rserif700, Rsans700, R700reg}, 700-extra = {Rserif700bot, Rserif700mid}, C360 = {Cserif360, Csans360}, C440 = {Cserif440, Csans440}, C540 = {Cserif540, Csans540}.

	Rserif150	Rsans150	Rserif360	Rsans360	Rserif440	Rsans440	Rserif540	Rsans540	Rserif700	R700reg	Rsans700	Rserif_700bot	Rsans700mid
Rserif150		-	-	-	-	-	-	-	[8.57]	-	-	[3.81 5.71 6.1	-
Rsans150			-	-	-	-	-	-	-	[14.28]	-	[6.19 6.66]	-
Rserif360				-	-	-	-	-	-	-	-	-	-
Rsans360					-	[2.86]	-	-	-	-	-	-	-
Rserif440						-	-	-	-	-	-	-	-
Rsans440							-	-	-	-	[2.86]	[2.86 3.33 3.8	[3.81]
Rserif540								-	-	-	-	[6.19 6.66]	-
Rsans540									-	-	-	-	-
Rserif700										-	-	-	-
R700reg											-	-	-
Rsans700												-	-
Rserif700bot													-
Rsans700mid7													

	R150	R360	R440	R540	R700	700-extra
R150		-	-	-	-	[3.81 4.28 4.76 5.24 5.71 6.19 6.66 7.14]
R360			-	-	-	-
R440				-	-	[3.81]
R540					-	-
R700						-
700-extra						

	Csans360	Csans440	Csans540	Cserif360	Cserif440	Cserif540
C360	-	-	[12.85]	-	-	-
Csans360		[12.85 13.33 13.8 1	[12.85 13.33 13.8]	-	[12.85 13.33]	[12.85 13.33 13.8]
Csans440			-	-	-	-
Csans540				-	-	-
Cserif360					-	-
Cserif440						-
Cserif540						

	C440	C540
C360	-	[12.38 12.85 13.33 13.8]
C440		-
C540		

Subtest 2: Statistical analysis

For subtest 2, the test in which all participants looked to the visual stimuli for a timespan that they determined themselves, some statistically significant differences were found: between Csans360_min and Cserif440 for the frequency bin 7.95Hz, and between Csans360_min and Cserif540_plus for the frequency bins 2.72Hz, 4.64Hz, 4.7Hz and 5.5Hz. The taller the font, the bigger the response.

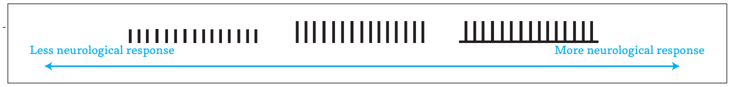
5.7. Test 3: conclusions

Taller stripe patterns result sometimes in more neurological response than lower stripe patterns. As a design researcher, I will not draw conclusions about what neurological response is going on and I am limiting myself to the differences in the visual stimuli.

Across the board, the taller the stripe/font, the more neurological response that occurred [Figure 12]. But there are three reasons why the hypotheses aren't proven. In subtest 1b for example, for the stripe patterns, the results were not found for all similar comparisons: for example, Rserif150 versus Rserif700 provides a significant difference, but Rsans150 versus Rsans700, or Rsans150 versus R700reg does not. The reason

FIGURE 12:

In the taller stripe patterns and in Serif_700bot, more neurological response was registered in a few frequencies (not in all).



therefore could not be determined based on this research. Second, there are more statistically non-significant comparisons than statistically significant comparisons [see Table 7]. This indicates that the effect is not that strong. Third, the results were even less consequent in the visual stimuli containing letters: for example, Csans 360 versus Csans 440 delivered a statistically significant result, but Cserif360 versus Cserif440 did not.

There was no statistically significant difference between any 700-font and the 700-extra fonts, indicating that a strike through will not result in a different neurological response than a 700-stripe pattern. However, it should be stated that there are more statistically significant differences between the lowest stripe patterns and the 700-extra fonts than with the 700-straight fonts. The results therefore suggest that the '700-extra' fonts cause more pronounced differences in the neurological response. A possible explanation for this is the number of details that the brain has to process.

Few statistically significant differences between sans and serif fonts were present, but none are consequent. The results of the three subtests differed. In subtest 1a, no statistically significant differences were found in the EEG's. This could be explained by the short time that the visual stimuli were shown (300ms). Subtest 1b did deliver results, probably because the time to look at the visual stimuli was long enough, 2,000ms.

6. Test 4: eye tracking while reading lines of text

6.1. Test 4: research question

1 Does a taller stripe pattern in text induce more eye closures/longer eye closures/a longer time to read compared to lower stripe patterns?

2 Does a more condensed and more unevenly spaced stripe pattern in text induce more eye closures/longer eye closures/a longer time to read compared to wider stripe patterns?

6.2. Test 4: research materials

The same testing materials as for test 3 were used, and measurements were done during the same three subtests [Figure 13].

FIGURE 13:

The testing materials for test 4.

6.3. Test 4: methodology

A total of 42 people participated. An eye tracker Eyelink 1000 Plus (RS Research, 2020) measured the number of times the eyes were closed and the time a participant needed to read all words set in each font. Data collection occurred at 500Hz.

6.4. Test 4: data collection

Test 2 was done at the same time as test 3, but results were not collected on subtest 1a. The reason is that the duration of 300ms is shorter than an eye closure can last, and that could provide false results.

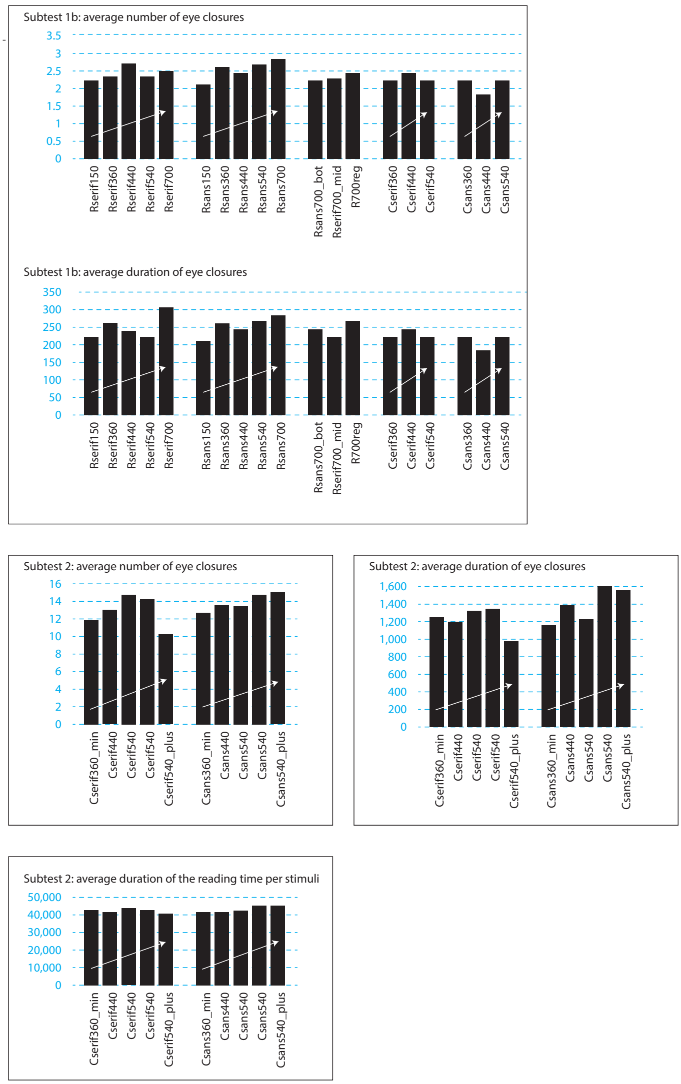
The data was filtered for the moments on which both eyes were closed at the same time. Closures shorter than 5/1000ms were left out, and if one of the eyes opened and closed again while the second eye remains closed, it was only counted as an additional eye opening if that first eye re-opened for a timespan longer than 100ms. This way, non-simultaneous closed eyes were corrected and vibrations in the eye lids were filtered out.

6.5. Test 4: results

With X as each of the ten fonts applied on a sentence, the results are calculated as (1) the average number of eyelid closures that occurred while reading font X, (2) the average duration of eyelid closures that occurred while reading font X. For test 1 was a fixed duration of 2,000ms used, but for test 2, the participant had control about the duration. Here, (3) the average time necessary to read font X was also taken into account. It is thus only present in Graph 'e'.

GRAPH 1 (a-e):

The results of the reading test. The white line throughout the graphs indicates the expected trend in the results.



6.6. Test 4: analyses

Mixed regression models were used with a random intercept per participant (which means that the information is correlated per person).

For the numbers of eye closures, test 1b is analyzed via a Poisson mixed regression and test 2 via a negative binomial mixed regression. There are no differences between the visual stimuli ($p>0.05$). Grouping visual stimuli together did not point to differences.

For the duration of the eye closures, test 1b is analyzed via a zero-inflated negative binomial mixed regression and test 2 via a negative binomial mixed regression. Statistically significant ($p<0.05$) results are:

- Eyes are statistically significantly longer closed on Rserif700 than on Rserif150 ($p=0.0479$).
- Eyes are statistically significantly longer closed on R700reg ($p=0.0409$), Rsans700 ($p=0.0254$) and Rsans540 ($p=0.0238$) than on Rsans150.

For the time a participant looked at the cues, test 2 is analyzed via a random effect model with a normal error division. None of the letterforms result in a statistically significantly longer reading time compared to other letterforms.

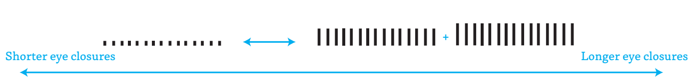
6.7. Test 4: conclusions

Regarding the stripe patterns, several of the higher fonts have a statistically significantly longer eye closure than the 150-fonts [Figure 14]. But it wasn't found in all of the tall stripe patterns. Additionally, the effect was not found in the fonts containing letterforms, so there is a difference between stripe patterns and letterforms.

While the 700-straight fonts resulted in statistically significant longer eye closures, the 'extra-700' fonts did not. That means the strikethrough affected the visual comfort of the stripe pattern when the effects on the eyes are measured.

The significant results only occurred when sans with sans was compared and serif with serif. There was no single statistically significant effect between horizontally differently spaced stripe patterns. This points out that even/unevenly spaced stripe patterns have less influence on the eye closures than lower/taller stripe patterns.

FIGURE 14:



Statistically significant results mostly occurred when comparing the extremes (150 vs 540 or 700).

7. General discussion

The role of stripe patterns in reading comfort

In all four tests, some statistically significant results were found. The height of the stipe pattern has several significant effects on the reading comfort, but effects differ between letters and stripe patterns. The horizontal distribution of the stripe pattern (distance between the stripes) had no useful significant effect in this study. Results of the current study do not directly point out if a serif or sans-serif typeface would be better for reading comfort.

The height of the stripe pattern

The results weren't always uniform. For example, the comparison of the serif pair 150 vs 700 versus the sans-serif pair of 150 vs 700 provided different results in the EEG. Also, not all stripe patterns resulted in the expected results. Therefore, some care is required when interpreting the results. Because there are only results on some places, they support the impression that taller stripe patterns have some effects, but it can 1) be too small to play a role within letters, 2) be too small to be measured by conventional devices or 3) be too small to hinder an audience with normal sensitivity to stressful visual stimuli. Therefore, it is too early to make solid conclusions. At the moment, no similar studies exist to compare with. It could be useful to do the same tests with people more sensitive to visual stress because excessive neurological response may hinder a sensitive audience more. We refer to people suffering migraine and others, such as mentioned in Aurora et al. (1991; mentioned in Wilkins, 2012). More research is needed.

No effects were found on the even/less even horizontal distribution of the stripe patterns. This contradicts other studies which did found effects of the stripe pattern in type design, such as Wilkins (2007) and Bessemans (2012). In both studies, making the spacing of the letters less even (e.g., a letter m without the equal spacing within) provided better reading results for certain target audiences. That this study did not provide statistically significant results probably depends on 1) that the horizontal spacing was not that distinguished in this study, 2) that there were differences in the methodologies and 3) that this study did not include children or poor readers.

The 700-extra's

It was expected that the '700-extra' group, the stripe patterns containing a strikethrough, would limit the visibility of stripe patterns. But contradictory, this group was ranked as least comfortable and caused a difference in neurological response compared to very low stripe patterns while the 700-group did not. While in the eye test, no significant results were measured on the 700-extra's. This suggests that a strikethrough makes it more difficult to process the image, but not to look at the image.

The same is noticed when comparing results of the stripe patterns with results of the letterforms. The lowest letterforms were not performing any better than the average height of the letterforms. This means that more design parameters than only the stripe pattern influence the reading comfort within letters.

I looked into which design parameter it could be and came to *the number of details per surface*. This parameter is expected to have an influence on reading comfort, because there is more information to process while reading. This parameter accompanies other parameters such as the height of the stripe pattern, how condensed a stripe pattern is, how (un)evenly spaced a stripe pattern is, and of course general parameters such as page layout, type size, etc. If this parameter would have an influence, this would also explain why the lowest fonts (360's) were ranked less comfortable than the average heighted fonts (440's): there were more details to process on a small surface.

No research was found that focused on this parameter within a typographic context. Future studies have to clarify the influence of this parameter.

Evaluation of the used approaches

In this study, four different tests applied four different methodologies. In all those tests, the results differ. For example, in the EEG-test, most of the statistically significant results are found in the 700-extra group, while in the eye-tracking test, most of the statistically significant results are found in the 700-straight group.

This does not indicate that any of the used methods were inappropriate, but it illustrates that different aspects of the reading process were measured in each test: the EEG evaluated neurological response, while the eye tracking focused on eye activity. So that will have been the reason for the differences in the results.

Those differences in the results point out that it would be difficult, if not impossible, to determine one parameter in which the reading comfort will be 'the best' under all circumstances. The overall reading comfort will probably depend on several design parameters at once and will depend on the condition on which the reading comfort is measured (neurological response, eye movements…).

Future studies

The results strongly suggest that more design parameters influence reading comfort simultaneously. Based on the results of this study, it can be stated that the design parameter 'height of the stripe pattern' probably plays a larger role than how condensed or how evenly spaced the stripe pattern is. But not all non-statistically significant results should be dismissed yet and the number of details per surface is a parameter about which no information is known yet. Those are to be further evaluated.

8. General conclusion

This study started from the general hypothesis that *"taller, less evenly spaced, and more condensed stripe patterns could be less effective for the reading comfort."* With this question, it strived to provide insights in making type design more comfortable to look at, and to evaluate the question if a serif or a sans-serif typeface would be better for reading. In this study, some effects were found on the tallest versus the lowest stripe patterns. But the results were not strong enough to prove the hypothesis right or wrong. Additionally, the stripe patterns on which most of the statistically significant effects were found are just out of range for regular type design.

This study strongly illustrates that: 1) results differed between the letterforms and the stripe patterns, showing that the reading comfort is not only influenced by the stripe pattern. Several results suggested that *'the number of design details per surface'* could be influential in the reading comfort, more than the height of the stripe pattern. 2) Results from the eye tracking and neurological responses were not identical. Someone studying reading comfort should take in mind that what is more comfortable for the eyes is not necessarily better for the neurological responses.

This study materials could not clarify if 1) letters already have the least hindering height of the stripe pattern [and that type designers thus intuitively do a good job], 2) if the testing devices were not refined enough. Because this was a preliminary study, the focus was not yet put on this kind of detail. Further studies are necessary to clarify.

9. Acknowledgements

The research was set up and performed by me, Maarten Renckens, on an independent basis. However, several people assisted. • I thank the Ethics Commission of the KULeuven (Catholic University Leuven) which reviewed and approved this study on November 2020 under number G-2020 11 2029. • I thank the members of the Flanders BCI Lab (https://www.flandersbcilab.be) for providing access to their equipment, Prof. Marc Van Hulle for overseeing the experiments, Dr. Benjamin

Wittevrongel and PhD candidate Axel Faes for sharing their technical expertise and their assistance in the data conversion. • I thank Krzysztof Basiński (assistant professor at the Medical University of Gdańsk, auditory cognitive neuroscientist and data analyst) for the analysis of the EEG-data in test one of this study. • I thank Dr. Johan Verbeeck (independent statistician) for the statistical analysis of the results of tests one, three and four.

References

Amazon. 2020. *"Modfans 3-pack leesbril mannen vrouwen leescomfort veerscharnier lichte rechthoekige heldere lens met tas +1.00 zwart" (MODFANS reading glasses man/woman reading comfort, light rectangle clear lens with bag +1.00 black).* [Online] https://www.amazon.nl/MODFANS-leesbril-leescomfort-veerscharnier-rechthoekige/dp/B07L8YRXTX [21 September 2020].

Bessemans, A. 2012. "*Letterontwerp voor kinderen met een visuele functiebeperking.*" [PhD dissertation], Leiden University & Hasselt University. http://hdl.handle.net/1887/20032.

British Independent Television Commission. 2001. *"Guidance note on flashing images and regular patterns in television."* [Re-issued by Ofcom: https://www.ofcom.org.uk/__data/assets/pdf_file/0023/104657/Section-2-Guidance-Notes.pdf]

Denksport. 2020. *"Zweeds Royaal – Editie 23 (Swedisch royal – edition 23)"*. [Online] https://www.denksport.be/beds-zweeds-ultra [21 September 2020].

De Standaard. 2006. "*Van Dale lanceert woordenboek voor slechtzienden (Van Dale launches dictionary for visual impaired)".* [Online] https://www.standaard.be/cnt/b342991060915 [21 September 2020].

Duin, M. 2014. *"21 tips voor meer leescomfort"* [Online] https://www.marjaduin.nl/21-tips-voor-meer-leescomfort [7 July 2020].

Gramfort, M.; Luessi, M.; Larson, E.; Engemann, D.a.; Strohmeier, D.; Brodbeck, C.; Goj, R.; Jas, M.; Brooks, T.; Parkkonen, L. & Hämäläinen, M.s. 2013. *"MEG and EEG data analysis with MNE-Python". Frontiers in Neuroscience, 7(267):* 1–13, 2013. http://doi:10.3389/fnins.2013.00267.

Jainta, S.; Jaschinski, W.; & Wilkins, A.j. 2010. *"Periodic letter strokes within a word affect fixation disparity during reading". Journal of Vision 10:* (13), 2.

Legge, G.e.; Bigelow, C.a. 2011. *"Does print size matter for reading? A review of findings from vision science and typography". Journal of Vision.* 11 (5). http://doi.org/10.1167/11.5.8.

Morrison, R.e.; & Rayner, K. 1981. *"Saccade size in reading depends upon character spaces and not visual angle". Perception & Psychophysics.* 30 (4): 395-396.

Nederlandse Taalunie. 2001. *"Woordenlijst Nederlandse taal (word list of the Dutch language)"*. Sdu Uitgevers. Den Haag/Antwerpen.

O'hare, L.; & Hibbard, P.b. 2011. *"Spatial frequency and visual discomfort". Vision Research 51:* 1767-1777.

O'hare, L.; & Hibbard, P.b. 2013. *"Visual discomfort and blur". Journal of Vision: April 2013, Vol.13, 7.* https://doi.org/10.1167/13.5.7.

Pennachio, O.; & Wilkins, A.j. 2015. *"Visual discomfort and the spatial distribution of Fourier energy". Vision Research 108:* 1-7.

Rayner, K. 1998. *"Eye movements in reading and information processing: 20 years of research". Psychological Bulletin.* 124 (3): 372-422.

Renckens, M. 2020. *"Consequently positioning the rhythm in type based on the letters' longest continuous black mass". Visible Language.* vol.54 (3): 32-47.

Sensotec. 2020. *"Wat is nieuw in SuperNova 18 (What is new in SuperNova 18)"*. [Online] https://sensotec.be/wat-is-nieuw-in-supernova-18 [21 September 2020].

Sr Research 2020. *"Specializing in eye tracking – Eyelink eye tracking – SR Research"*. [Online] https://www.sr-research.com [3 September 2020].

Van Dale. 2015. *"Van Dale Groot woordenboek van de Nederlandse taal (Van Dale great dictionary of the Dutch language)"*. Van Dale Uitgevers. Utrecht/Antwerp. ISBN: 9789460772221.

Vivlio. 2020. *"Wat zijn de verschillende e-readers modellen van Vivlio? (What are the different e-reader models of Vivlio?)"*. [Online] https://help.vivlio.com/hc/nl/articles/360003847020-Wat-zijn-de-verschillende-e-readers-modellen-van-Vivlio- [21 September 2020].

Wilkins, A.; Nimmo-Smith, I.; Tait, A.; Mcmanus, C.; Della Sala, S.; Tilley, K.a.; Barrie, M.; & Scott, S. 1984. *"A neurological basis for visual discomfort"*. Brain. 107: 989-1017.

Wilkins, A. 2012. *"Origins of Visual Stress".* In: *Visual aspects of dyslexia".* Oxford: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199589814.003.0004.

Wilkins, A.; & Nimmo-Smith, M.i. 1987. *"The clarity and comfort of printed text". Ergonomics.* 30 (12): 1705-1720.

Wilkins, A.j.; Smith, J.; Willison, C.k. Baere, T.; Boyd, A.; Hardy, G. Mell, L.; Peach, C.; & Harper, S. 2007. *"Stripes within words affect reading". Perception.* 36: 1788-1803.

Author

Maarten Renckens Independent Multidisciplinary Researcher

info@maartenrenckens.com

Maarten Renckens is a multidisciplinary person. He combines functions as design researcher (designer & Linotype technician) / landlord / Environment Expert (urban planning & building permissions). But above all, he does the things he likes to do. As someone who used to have light reading difficulties himself, he took the task to enlarge the knowledge about the design parameters that could influence the reading comfort of texts. Within this context, he tests different design parameters with different audiences create more awareness of proper text design.