The usage of colours as a text highlighting technique

Paweł Krzos

Wrocław University of Science and Technology

Faculty of Computer Science and Management

Abstract: The need of using text highlighting techniques exists in various cases. There are numerous different techniques available, among others: font colour, background colour, underlining, font weight, font style or capitalization. The guidelines describing the usage of colours in different contexts can be found in many sources. Over the last century studies in this field have been conducted but the impact of colours on the pop-out effect has not been comprehensively investigated so far. The present study analyzes that phenomenon and establishes preliminary ranking of best colours to use for text highlighting basing on opinions of 82 students. Two highlighting methods: text colour and background colour, with six basic colours: red, green, blue, cyan, magenta and yellow, have been examined. The results show that significant differences between colours exist. The stability of answers has been confirmed but some of the data is inconsistent. Limitations as well as directions for future work are also described.

Key words: text highlighting, colours usability, visual document analytics

1. Introduction

1.1. Text highlighting

The very most of information is presented as a natural language text. Nowadays we are struggling with huge amount of data which leads the readers to skimming the documents (Nielson, 1997). Because of that it is very important to attract attention to the most essential ideas. It can be done especially by text highlighting. The best known methods to distinct a part of a text are: background colour, font colour, font size, font style (italics, subscript), font weight (bold font), font family, underlining, additional borders, letter spacing, shadowing, or capitalization. Some work has already been done to understand the usage of highlighting in paper and digital documents (Churchill, 2000; Marshall, 1997; Schilit, 1998). It has also been proved that highlighting has remarkable meaning in the field of educational psychology (Peterson, 1992; Nist, 1987; Silvers, 1997). Some articles had been writ-

Correspondence to:
Paweł Krzos
Politechnika Wrocławska
Wydział Informatyki i Zarządzania
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
Tel.: +48 793 777 204
E-mail: pawel.krzos@pwr.edu.pl

ten before the computer displays started to be used but they are also important to understand the flow of studies in the area of visual analytics. This article is focused on the context of usability of contemporary software and deepens Strobelt et al. (2016) analysis. Strobelt has conducted an experiment where the following aspects have been determined:

– a ranking of nine commonly-used text highlighting techniques,
– the degree of visual interference between pairs of highlighting techniques,
– the effectiveness of techniques for visual conjunctive search.

It was an in-depth study, however only one colour has been used for font and background highlighting. The authors suggested developing next studies to check if there would be differences between other colour combinations. This paper aims to verify it.

2. Colour in text

At first some basic definitions should be explained. Readability is the ease with which a reader can understand a written text. Legibility has an influence on readability and describes the ease of distinguishing individual letters or characters from each other.

There are two main colour models: subtractive and additive. The first one explains the mixing of a limited set of dyes which enables obtaining a wider range of colours. The most popular subtractive model, used especially for printing purposes, is CMY (Hasan, 2012) with cyan, magenta and yellow colours. The second one explains the mixing of a number of different beams, the most often: red, green and blue (RGB) (Hasan, 2012), which altogether create white colour (instead of multiplication to black in CMY model). The same colours can be described by various models where besides hue also other visual properties such as luminance are taken into consideration. This study is limited only to examine the 6 basic colours coming from subtractive and additive models.

The first research in the field of using colours has been carried out by Le Courier over a hundred years ago (Livre, 1912). As a result the ranking of legibility depending on font and background colour has been developed. The most legible pair was black and yellow what is quite surprising in comparison to widely used combination of black and white. The study has been revised in 2008 by Humar et al. (2008), who tested the legibility of a web page text presented on CRT displays. They have proved that negative polarity (pairs yellow/ black and white/ blue) performs best in their case. The comparison of both studies is presented in Table 1 below.

Table 1. The text legibility ranking of Humar et al. and Le Courier studies

Text/ background	Study	
	Humar et al.	Le Courier
Yellow/ black	1	7
White/ blue	2	5
Black/ yellow	3	1
White/ black	4	10
Black/ white	5	6
Blue/ white	6	4
Red/ yellow	7	11

Text/ background	Study	
	Humar et al.	Le Courier
White/ red	8	8
Red/ white	9	3
Red/ green	10	13
Green/ white	11	2
White/ green	12	9
Green/ red	13	12

S o u r c e: Author's own elaboration based on Humar, 2008.

Both researches have concentrated on the text legibility in general. They have not investigated the impact of colour on the highlighting techniques though. Over the next century since the first Le Courier's experiment a number of studies have been carried out. A substantial insight has been given in a series of surveys by M. Tinker and D. Paterson (Tinker, 1928; 1929a; 1929b; 1929c; 1929d; 1929e; 1932; 1940; 1942; 1944; 1946; 1955; 1963). L. Matthews has proved that hue, in contrast to luminance, has no statistically significant effect on the readers' visual performance (Matthews, 1987). The similar results have been achieved by S. Pastor (1990). The topic has been further investigated, among others by Hill and Scharff (Hill, 1997), van Schaik (Ling, 2002; Pearson, 2003), Wang and Chen (Wang, 2003), Hall and Hanna (Hall, 2004) and Humar et al. (Humar, 2008). Summarizing, many various studies have been carried out but the consistent conclusion cannot be drawn. The results are often contradictory to each other and it would be nearly impossible to establish only one colour ranking. Despite of that Strobelt et al. (Strobelt, 2016) have proven that there are statistically significant differences in search performance between various highlighting techniques. They have also suggested checking if other colour combinations would have any influence on the results. Therefore, the main objective for this study is to verify the hypothesis that the preferences in choosing the colour of text highlighting exist.

3. Method

Six basic colours have been chosen for testing: cyan (#00FFFF), magenta (#FF00FF), yellow (#FFFF00), red (#FF0000), green (#00FF00) and blue (#0000FF). The background colour for whole document has been fixed to white and the font colour for background highlighting has been fixed to black. The two types of highlighting: font colour and the background colour, have been investigated in separate surveys. Each combination of two colours has been shown to participants, 15 pairs for font colour and 15 pairs for background colour in total. Each pair has been shown per 10 seconds and the participant had to fill the attached questionnaire answering the question: 'Which colour would you use for highlighting the text?'. One word located exactly in the centre of the paragraph has been highlighted. All colour variants have been displayed to participants before the survey so that they should be well informed about all possibilities to choose from. 82 participants have been recruited as volunteers (gender: 43 females, 39 males, age 17–27). The pairs have been shown in random order. The col-

our blind test has been carried out to exclude colour blind people from the survey, yet everybody has passed it. The survey has been carried out twice with one hour break to check the stability of answers. The figures below depict the exemplary pairs which have been shown to the participants. The content has been presented on the 17″ LCD display. The participants have been sitting 1 metre in front of the screen. There were no distracters.

The results have been counted and normalized to the scale of 1–9 so that they could be an input to the AHP matrix (Saaty, 1977). The significance of differences between colours has been checked using ANOVA and also pairwise comparison with student's t-test. The stability of ranking between two runs has been checked with dependent t-test for paired samples. The consistency of ranking has been verified with AHP. No special statistical software has been used. The calculations have been done in Microsoft Excel.

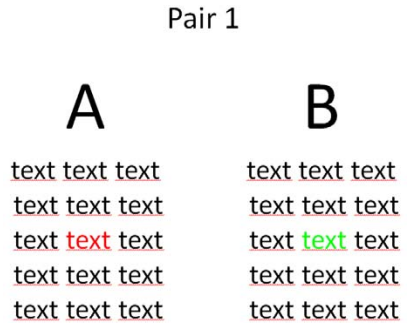

Figure 1. The exemplary pair from the first survey: highlighted fonts (left: red, right: green)

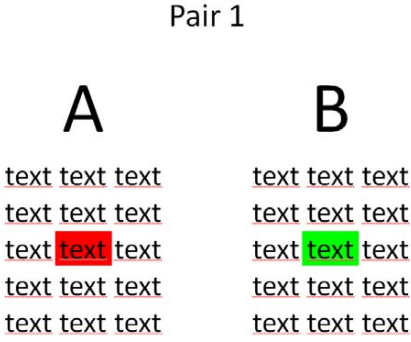

Figure 2. The exemplary pair from the second survey: highlighted backgrounds (left: red, right: green)

4. Results and discussion

4.1. Study 1: font colour

The results of the study are presented in the tables below. In Table 2 colours have been grouped according to the results of the significance test for each pair (results from the second run are shown in brackets). Tables 3 and 4 present AHP matrixes with coefficients calculated as normalized count of answers for the first and the second run correspondingly.

Table 2. The ranking of font colours

Colour	G	R	M	B	C	Y	Mean	StDev
G	A (A)						0.74 (0.78)	0.23 (0.20)
R		B (B)	B (B)	B (B)			0.57 (0.57)	0.31 (0.32)
M		B (B)	B (B)	B (B)			0.56 (0.54)	0.22 (0.25)
B		B (B)	B (B)	B (B/C)	(C)		0.54 (0.49)	0.26 (0.30)
C				(C)	C (C)		0.38 (0.41)	0.30 (0.30)
Y						D (D)	0.22 (0.20)	0.26 (0.26)

S o u r c e: Author's own elaboration.

Table 3. AHP matrix for font colours, the first run

Colour	R	G	B	C	M	Y
R	1.00	0.36	2.95	2.17	0.84	5.10
G	2.76	1.00	4.70	5.49	4.12	6.85
B	0.34	0.21	1.00	4.51	1.78	5.29
C	0.46	0.18	0.22	1.00	0.23	3.93
M	1.20	0.24	0.56	4.32	1.00	6.46
Y	0.20	0.15	0.19	0.25	0.15	1.00

S o u r c e: Author's own elaboration.

Table 4. AHP matrix for font colours, the second run

Colour	R	G	B	C	M	Y
R	1.00	0.32	3.15	1.78	1.00	5.49
G	3.15	1.00	5.88	6.85	3.93	7.63
B	0.32	0.17	1.00	2.76	1.39	5.29
C	0.56	0.15	0.36	1.00	0.56	3.54
M	1.00	0.25	0.72	1.78	1.00	6.66
Y	0.18	0.13	0.19	0.28	0.15	1.00

S o u r c e: Author's own elaboration.

The order of colours is the same for both runs of the test. ANOVA showed significant differences between the colours (first run: F=37.03, second run: F=38.80). The best colour for text highlighting is green while yellow is the worst. The rest, apart from cyan, is quite good without meaningful differences between each other. Cyan has been higher rated in the second

run which led to weaker difference in ranking from blue, no longer statistically significant. Dependent t-test for paired samples has proved that there are no significant differences between the runs (the biggest Z-score: 0.89). In both cases CR values are acceptable (first run: CR=0.054, second run: CR=0.020) which means that the results are consistent.

The interesting finding is that green has been chosen as the most preferable colour. In this study any reading context has been intentionally removed, however there are many examples of connection between font colour and the type of highlighted text (e.g. blue hyperlinks). Despite of this, green has been the mostly selected colour. The first three colours are compliant with the ranking established by Le Courier (cf. Table 1). The low rating of yellow can be explained by the poor contrast with white background so that the text is hardly readable.

4.2. Study 2: background colour

The results are presented just as those for Study 1.

Table 5. The ranking of background colours

Colour	C	Y	G	M	R	B	Mean	StDev
C	A (A)	A (A)	 (A)				0.82 (0.80)	0.22 (0.20)
Y	A (A)	A/A2 (A)	A2 (A)				0.80 (0.75)	0.18 (0.21)
G	 (A)	A2 (A)	A2 (A)				0.70 (0.75)	0.17 (0.19)
M				B (B)	B (B)		0.35 (0.35)	0.15 (0.15)
R				B (B)	B (B)		0.30 (0.31)	0.17 (0.18)
B						C (C)	0.03 (0.04)	0.10 (0.14)

S o u r c e: Author's own elaboration.

Table 6. AHP matrix for background colours, the first run

Colour	R	G	B	C	M	Y
R	1.00	0.13	8.80	0.12	0.34	0.12
G	7.83	1.00	8.61	0.28	8.02	0.27
B	0.11	0.12	1.00	0.12	0.12	0.11
C	8.22	3.54	8.41	1.00	8.22	8.80
M	2.95	0.12	8.02	0.12	1.00	0.12
Y	8.41	3.73	8.80	0.11	8.02	1.00

S o u r c e: Author's own elaboration.

Table 7. AHP matrix for background colours, the second run

Colour	R	G	B	C	M	Y
R	1.00	0.13	8.41	0.12	0.39	0.13
G	7.83	1.00	8.80	0.46	7.63	0.72
B	0.12	0.11	1.00	0.12	0.13	0.12
C	8.22	2.17	8.61	1.00	7.83	1.98
M	2.56	0.13	7.83	0.13	1.00	0.12
Y	7,80	0,38	7,80	0,26	8,20	1,00

S o u r c e: Author's own elaboration.

The same as in Study 1 the order of colours has not changed in two runs. ANOVA showed also that significant differences exist (first run: 298.7, second run: F=239.5). Cyan has been evaluated as the best colour but green and yellow are just behind. They are followed by significantly different group of magenta and red colour while blue is located at the very end of the ranking. No significant differences between the runs have been found (the biggest Z-score: 1.10). The AHP analysis showed that the results from the first run are strongly inconsistent (first run: CR=0.235, second run: CR=0.087). The same as in Study 1, the second run had more coherent answers but they are still near the border of tolerance.

The opinions about background colours have been expressed more strongly (many values in AHP matrixes are close to extremes). Though, the results are less consistent than those from the first study. Pastel colours (cyan, green, yellow, magenta) are well known to be used in highlighters. There is no meaningful difference between top three of them. Magenta has in total lower score, probably because it is perceived as females' colour and is not willingly used by males. Red and blue perform well in font colour distinction but they are definitively not a good choice for background highlighting. Especially blue has been evaluated very poorly. It can be explained by low contrast between black text and dark background. It would be an interesting study to check how dark background colours generate a pop-out effect with other font colours (e.g. white or yellow). It remains, however, beyond the scope of this paper.

5. Limitations and future work

The survey has been carried out under several limitations. Only 6 colours have been taken into consideration. The main settings (surrounding text and background, font colour in the second survey) have been set to fixed values. The main factor differentiating the colours was only hue. The context of reading has been consciously taken out. In future other colour combinations can be investigated. Developing the tests under conditions closer to reality (authentic documents or web pages) would also be a good idea. The study can be revised on bigger data sample as well as the differences between various groups (gender, age etc.) can be examined. The relationship between general preferences (e.g. somebody's favourite colour) and colour ranking can also be investigated. Despite of the same order of colours in two runs of each study the issues with consistency of the ranking arisen. The insufficient preparation of participants can explain the differences between both runs but this topic should be

further investigated, e.g. more runs can be executed. The other methodology can be used to obtain more objective results. The test environment similar to once proposed by Strobelt et al. (2016) can be prepared or even the eye-tracking devices can be used.

6. Conclusions

This study confirms that the differences between willingness of using colours as a highlighting technique exist. The ranking of font colours and background colours has been established and has not changed in both runs of the experiment (G/R/M/B/C/Y and C/Y/G/M/R/B correspondingly). The issue with the consistency of given answers is present though. The survey gives a preliminary insight and has potential for further development. Despite of all limitations it can be clearly seen that incorrect usage of colour can be easily noticed by readers and it has an impact on the text legibility.

Acknowledgements

Special thanks to Professors Jerzy Grobelny and Rafał Michalski from the Wrocław University of Science and Technology for valuable comments and suggestions. The acknowledgements are given also to all participants for their commitment in this study.

References

Churchill, E. T., Trevor, J., Bly, S., Nelson, L., Cubranic, D. (2000). Anchored conversations: chatting in the context of a document. In: *Proceeding of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 454–461). New York: ACM Press. ISBN 1581132166.

Hall, R. H., Hanna, P. (2004). The impact of web page text-background colour combinations on readability, retention, aesthetics and behavioural intention. *Behaviour & Information Technology*, *23* (3), 183–195.

Hill, L. S. (1997). Readability of screen displays with various foreground. Background color combinations, font types and word styles. In: *Proceedings of the Eleventh National Conference on Undergraduate Research* (pp. 742–746). Austin, Texas.

Humar, I., Gradišar, M., Turk, T. (2008). The impact of color combinations on the legibility of a Web page text presented on CRT displays. *International Journal of Industrial Ergonomics, 38* (11–12), 885–899.

Ibraheem, N. A., Hasan, M. I., Khan, R. Z., Mishra, P. K. (2012). Understanding color models: A review. *ARPN Journal of Science and Technology*, *2* (3), 265–275.

Ling, J., van Schaik, P. (2002). The effect of text and background colour on visual search of web pages. *Displays*, *23*, 223–230.

Livre, L. C. (1912). Lisibilité des affiches en couleurs. In: *Cosmos* (p. 255). Sheldons Limited House.

Marshall, C. (1997). Annotation: from paper books to the digital library. In: *Digital Libraries '97 2nd ACM International Conference on Digital Libraries* (pp. 131–140). New York: ACM Press. ISBN 0897918681.

Matthews, L. M., Mertins, K. (1987). The influence of color on visual search and subjective discomfort using CRT display. In: *Proceedings of the Human Actors Society 31st Annual Meeting* (pp. 1271–1275). Santa Monica: HFES.

Nielsen, J. (1997). *How users read on the Web* [online, accessed: 2017-08-10]. Retrieved from: http://www.useit.com/alertbox/9710a.html.

Nist, S. H., Hogrebe, M. C. (1987). The role of underlining and annotating in remembering textual information. *Reading Research and Instruction*, *27* (1), 12–25.

Pastoor, S. (1990). Legibility and subjective preference for color combinations in text. *Human Factors*, *32* (2), 157–171.

Pearson, R., van Schaik, P. (2003). The effect of spatial layout of and link colour in web pages on performance in a visual search task and an interactive search task. *International Journal on Human-Computer Studies*, *59* (3), 327–353.

Peterson, S. (1992). The cognitive functions of underlining as a study technique. *Reading Research and Instruction*, *31* (2), 49–56.
Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, *15* (3), 234–281.
Schilit, B. G. (1998). Beyond paper: Supporting active reading. In: *Proceedings of CHI '98* (pp. 249–256). New York: ACM Press.
Silvers, V. L., Kreiner, D. S. (1997). The effects of pre-existing inappropriate highlighting on reading comprehension. *Reading Research and Instruction*, *36* (3), 217–223.
Strobelt, H., Oelke, D., Kwon, B. C., Schreck, T., Pfister, H. (2016). Guidelines for effective usage of text highlighting techniques. *IEEE Transactions on Visualization and Computer Graphics*, *22* (1), 489–498.
Tinker, M. P. (1928). Influence of type form on speed of reading. *Journal of Applied Psychology*, *12*, 359–368.
Tinker, M. P. (1929a). Studies of typographical factors influencing speed of reading: II. Size of type. *Journal of Applied Psychology*, *13* (2), 120–130.
Tinker, M. P. (1929b). Studies of typographical factors influencing speed of reading. III. Length of line. *Journal of Applied Psychology*, *13* (3), 205–219.
Tinker, M. P. (1929c). Studies of typographical factors influencing speed of reading: V. Simultaneous variations of type site and line length. *Journal of Applied Psychology*, *13* (2), 72–78.
Tinker, M. P. (1929d). Studies of typographical factors influencing speed of reading: VI. Black type versus white type. *Journal of Applied Psychology*, *13* (2), 241–247.
Tinker, M. P. (1929e). Studies of typographical factors influencing speed of reading. VII. Variations in color of print and background. *Journal of Applied Psychology*, *13* (2), 471–479.
Tinker, M. P. (1932). Studies of typographical factors influencing speed of reading: X. Style of type face. *Journal of Applied Psychology*, *16*, 605–613.
Tinker, M. P. (1940). *How to Make Type Readable: A Manual for Typographers, Printers, and Advertisers, Based on Twelve Years of Research Involving Speed of Reading Tests Given to 33,031 Persons*. New York: Harper & Brothers.
Tinker, M. P. (1942). Reader preferences and typography. *Journal of Applied Psychology*, *26*, 38–40.
Tinker, M. P. (1944). Eye movements in reading black print on white background and red print on dark green background. *American Journal of Psychology*, *57*, 93–94.
Tinker, M. P. (1946). Readability of mixed type forms. *Journal of Applied Psychology*, *30*, 631–637.
Tinker, M. P. (1955). *Tinker speed of reading test*. Minneapolis: University of Minnesota Press.
Tinker, M. P. (1963). *Legibility of Print*. Ames: Iowa State University Press.
Wang, H. C. (2003). Effects of screen type, Chinese typography, text/background color combination, speed, and jump length for VDT leading display on users' reading performance. *International Journal of Industrial Ergonomics*, *31* (4), 249–261.

Wykorzystanie kolorów jako techniki wyróżniania tekstu

Abstrakt: W różnych sytuacjach występuje potrzeba wyróżnienia fragmentu tekstu. W tym celu dostępnych jest wiele różnych technik, między innymi: kolor czcionki, kolor tła, podkreślenie, wielkość i styl czcionki, kapitalizacja. Wytyczne opisujące użycie kolorów w różnych kontekstach można odnaleźć w wielu źródłach. W ciągu ubiegłego stulecia przeprowadzono liczne badania w tej dziedzinie, jednak jak dotąd nie badano wpływu kolorów na intensywność efektu wyróżniania (ang. *pop-out effect*). Na podstawie przeprowadzonej ankiety i analizy tego zjawiska określony został ranking najlepszych kolorów używanych do wyróżniania tekstu. Zweryfikowane zostały dwie metody (kolor czcionki i kolor tła) dla sześciu podstawowych kolorów. Badanie zostało przeprowadzone dwukrotnie na tej samej grupie respondentów, aby sprawdzić stabilność odpowiedzi. Spójność opinii została zbadana przy użyciu metody AHP. Wyniki wskazują, że istnieją

znaczne różnice między kolorami. Preferencje zostały pogrupowane w klasy. Stabilność odpowiedzi została potwierdzona, jednak niektóre wyrażane poglądy nie są spójne. Przeprowadzone badanie stanowi wstęp do omawianego tematu. Wiele czynników pozostaje poza zakresem niniejszego artykułu i stanowi potencjał do dalszych eksperymentów (m.in. specyfika kulturowa, odcienie kolorów, kontekst prezentacji).

Słowa kluczowe: wyróżnianie tekstu, kolory w tekście, wizualna analiza dokumentu