ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20160225140202

TEXT READABILITY AND LEGIBILITY ON iPad WITH COMPARISON TO PAPER AND COMPUTER SCREEN

Darijo Čerepinko, Damira Keček, Mario Periša

Original scientific paper
In this paper authors examine readability and legibility of text on iPad and compare it partially to reading from computer screen and paper. Following previous research, two independent variables were measured: speed of reading (T) and subjective difficulty of reading (Z). 220 university students aged 18÷48 participated in the experiment and were divided into 11 groups of 20 participants. Key findings are as follows: Sans serif Gotham font is more readable and legible on iPad than Minion Pro serif font, although, when presented in bigger font sizes there is no significant difference between the two letter cases; two column spread of text was the hardest to read, giving fewer number of characters per row; layout of text with 79 characters per row and above should be readable and legible enough for reading texts on iPad; considering quality of reading, texts displayed on iPad can replace texts printed on paper.

Keywords: *iPad; legibility; readability; text*

Čitkost i čitljivost teksta na iPadu i usporedba s čitanjem s papira i zaslona računala

Izvorni znanstveni rad
U radu autori istražuju čitkost teksta na iPad uređaju i djelomično ga uspoređuju s čitanjem sa zaslona računala i papira. Na temelju ranijih istraživanja, mjerene su dvije nezavisne varijable: brzina čitanja (T) i subjektivna težina čitanja (Z). U istraživanju je sudjelovalo 220 studenata starih između 18 i 48 godina podijeljeno u 11 grupa od po 20 sudionika. Ključni rezultati su: Sans serif font Gotham čitkiji je i čitljiviji na iPadu od Minion Pro serifnog fonta, iako kad su veličine fonta veće, nema statistički značajne razlike između njih; zbog manjeg broja znakova po retku, dvostupačni prijelom najteže je čitljiv; tekst prelomljen u minimalno 79 znakova po retku i više dovoljno je čitak za čitanje na iPadu; s obzirom na kvalitetu čitanja, tekstovi prikazani na iPadu mogu zamijeniti tekstove tiskane na papiru.

Ključne riječi: *iPad; čitkost; čitljivost; tekst*

1 Introduction

With introduction of new reading technologies such as computer screens, mobile phone and tablets, researchers find it essential to analyse how the changes in the presentation mode and/or in medium affect readers' ability to read different texts. Most of the research is focused either on reading speed or reading comprehension of the readers answering how different presentation modes influence readability and legibility of text.

Reading process is considered to be the ability of the reader to recognize and identify visual forms of language – written words and letters [e.g. 1, 2, 5] that are organized by the means of typography such as, among others, typeface or font size [3, 4]. Studies of reading go as far as into the 19th century. Over the past two hundred years there were numbers of studies employing many different methods and exploring various aspects of texts in print and on screen but with three significant limitations: 1) there is extremely limited number of papers dealing with readability and legibility of text displayed on tablet computers and there is almost no guidelines for layout of applications, user interfaces and/or documents; 2) almost all of the research papers were presenting findings characteristic for English language texts and readers; and 3) there is very limited number of studies that compare readability and legibility of text between three different display technologies [5][1].

Following those limitations, in this paper we present the results of comparative analysis of reading of the text in three different presentation modes – on paper, computer screen, and iPad, with a more detailed analysis of readability of text on iPad devices. For the experiment we used two different typefaces – serif and sans serif; four different font sizes – 10, 12, 14 and 16 pt; and single and two column spread. Language of the text is also different than English (i.e. Croatian) and might be used for further comparisons and research.

2 Overview of the previous research

Some of the earliest papers on the matter, as shown by Dillon [7], indicated there were significant differences in results between paper vs. screen considering various performance outcomes such as reading speed, accuracy and comprehension. Based on his comprehensive comparative review of empirical literature, Dillon concluded that reading from screen could be up to 30 percent slower than reading from paper. Although this finding is supported by the majority of the work presented in his review, there were some papers that have found no significant difference when the text is read in the same or similar circumstances [e.g. 8, 9].

Fifteen years after Dillon, Noyes and Garland [10] conducted another critical overview of the relevant literature, suggesting in their findings that overall results are again inconsistent, although with somewhat greater equivalence than in Dillon's overview. They also indicated a new, more sophisticated set of measures and new methodology used for comparison between the paper and screen that are more task oriented as opposed to traditional outcomes, such as reading speed, accuracy or fatigue. These new methods include non-standardized open-ended tasks, non-standardized closed tasks, and

[1] Almost as an exemption to the rule, Nielsen (2010) conducted reading speed study using four different display technologies – paper, computer screen, Kindle e-reader and iPad.

standardized tasks that were administered to participants. However, one of the problems of such an approach is that different technologies apply different performance demands and comparison between them could be often difficult or even impossible. Partially because of that, there are still a number of researches conducted by using traditional measurements, especially when concerned with the issues of spatial presentation of text and typography [e.g. 11, 12, 13, 14].

Much research compares what effect varying font size, typeface or other typographic elements induce on readers, mainly comparing differences in one display technology (i.e. either paper or on screen) and then comparing results between different technologies [see 5]. Sheedy et al. [15] have found, for example, that sans serif fonts like Ariel and Verdana are more legible on screen than serif fonts such as Times New Roman and that lowercase letters are up to 20 percent more legible than lowercase words when read on screen. They also define optimal font size of 9 px (around 10 pt) which contradicts findings relating reading from paper where serif fonts are more legible and optimal font size is around 11 pt. Chandler [16] finds optimal reading font size around 12 pt and Subbaram at 14 pt [17]. Legibility of text is also the subject of spacing and leading, which can be observed trough some rules of Gestalt theory [3] and, according to the Theory of reading, visual span or the number of letters/signs that can be seen in single eye fixation [18]. Some researchers point out the importance of previous experience of readers stating that readability and legibility are not something that is inherent to typeface 'by nature' but is, in part, the result of readers' familiarity with the typeface [e.g. 19].

A study from 2010 conducted by Nielsen [20] on 32 participants measured reading speed and user satisfaction trough four display technologies: paper, computer screen (PC), Kindle e-reader, and iPad. Their findings (although lacking some technical details about computer screen and paper printout) found that reading the same text from iPad1 is 6,2 % slower than from paper. Reading from Kindle2 e-reader is 10,7 % slower than paper and there is no significant difference between two compared to each other. Reading speed for computer screen was not included into paper. Concerning user satisfaction, they found that, 'on a 1÷7 scale, with 7 being the best score [.]iPad, Kindle, and the printed book all scored fairly high at 5,8, 5,7, and 5,6, respectively. The PC, however, scored an abysmal 3,6' [20] concluding that 'this study is promising for the future of e-readers and tablet computers'.

3 Research methods

Following the Nielsen [20] research and methods used in similar studies described by Tinker [3], Tarasov et al. [6] and Dillon [7], two independent variables were measured: speed of reading (T) and subjective difficulty of reading (Z). Subjective difficulty is subjective feeling each participant had while reading the administered test piece, compared to their usual reading experience and could be seen similar to Nielsen's satisfaction (but opposite in polarity) and to some extent linked to fatigue variable measured in earlier research.

220 university students aged 18÷48, divided into 11 groups of 20 participants, were asked to read short excerpt (169 words; 853 characters) from an actual article published in Croatian daily newspaper "Jutarnji list" with medium reading difficulty (*LIX* = 37,35), calculated by using adapted LIX readability formula for Croatian language [21]. Text was presented in 10 different test pieces, as shown in Tab. 1, with layout elements defined through a series of interviews and focus group with professionals in the field of graphic and web design [22] prior to the experiment.

The test pieces contained two distinct typefaces frequently used in digital formats: Minion Pro (serif typeface) and Gotham (san serif typeface), four different font sizes (10, 12, 14 and 16 pt) and one column or two column layout. Average number of characters in a row was from 79 (16 px) up to 112 in one column spread and from 42 to 49 in two column spread. Test piece printed on paper was printed in black on white with printing resolution of 300 dpi. Test piece presented on computer screen was presented on a computer screen with WXGA 1366 × 768 px resolution. Test pieces presented on iPad were presented on a screen of 1024 × 768 px and 132 ppi resolution.

Table 1 Test pieces characteristics

Test Piece/Layout	Display technology	Typeface	Font size	No of Columns
Test piece 1	paper	Minion Pro	12	1
Test piece 2	computer screen	Minion Pro	12	1
Test piece 3	iPad	Minion Pro	12	1
Test piece 4	iPad	Gotham	12	1
Test piece 5	iPad	Minion Pro	12	2
Test piece 6	iPad	Minion Pro	10	1
Test piece 7	iPad	Minion Pro	14	1
Test piece 8	iPad	Minion Pro	16	1
Test piece 9	iPad	Gotham	14	1
Test piece 10	iPad	Gotham	16	1
Test piece 11	iPad	Minion Pro	12	1

Groups 3 and 11 had exactly the same task – to read the text on the same device, written in the same typeface and font size – with the exemption that Group 11 participants were administered with a short series of questions concerning the text content in order to get the confirmation that they actually read the article. Their results were later tested with appropriate statistical tests and showed no significant differences compared to the results of the participants in Group 3. Subjective difficulty of reading (Z) was measured by using a post-reading survey with one question Likert questionnaire scaling from 1 – Extremely difficult to 10 – Extremely easy.

4 Results and discussion

Results of the reading test are shown in Tab. 2. Values in columns (T1 to T11 and Z1 to Z11) are presenting the results of reading speed (T) and subjective difficulty of reading (Z) for each participant (G1 to G20). Due to its visual simplicity [23], the results of the statistical analysis of data are presented visually in a form of Box-and-Whisker Plots shown in Fig. 1 and 2.

As explained above, since the participants in the Group 11 were used as the control group and their results were used for comparison against the results of an actual

test group (Group 3), their results are not included in the statistical analysis of data shown below. Since there was no statistically significant difference between the two groups, results from Group 11 are not further analyzed.

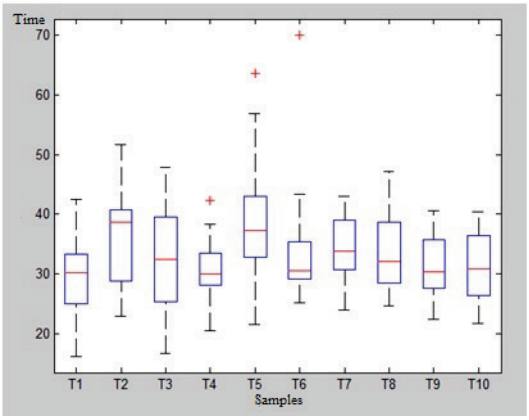

Figure 1 Reading speed (T) results across the groups 1 to 10

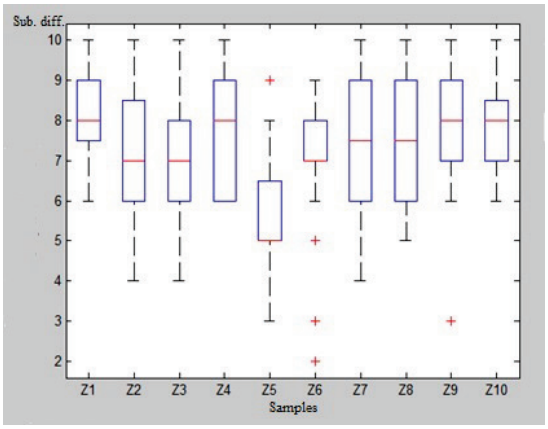

Figure 2 Subjective difficulty of reading (Z) results across the groups 1 to 10

Red line in plots represents the median value for the observed sample. Rectangle represents 50 % of all the data, ranging between 25 to 75 % of the data for the observed sample. Upper and lower horizontal bars represent lowest and highest data values in the sample, with the exception of the outliers which are presented with a red plus sign. Since measured results differ from normal distribution, tested and confirmed by Kolmogorov–Smirnov nonparametric test, for the purpose of testing the null hypothesis the Kruskal-Wallis test is used and the results are presented in Tab. 3.

As shown in Fig. 1 and 2 and Tab. 3, the shortest average reading time is determined for sample T1 and the longest average reading time is determined for sample T5. Higher values of the median are determined for samples T2 and T5 whereas the lowest median value is determined for sample T4. In all the samples median value corresponds quite highly with the mean value. As shown in Tab. 3, there is statistically significant difference between the samples considering the speed of reading (p = 0,0057). Pairs of samples that are differing significantly are (T1, T2), (T1, T5), (T1, T7), (T2, T4), (T2, T9), (T2, T10), (T4, T5), (T5, T6), (T5, T9) and (T5, T10). Accordingly, for the reading difficulty (Z), the highest average value is determined for the sample Z1 and the lowest average value is determined for the sample Z5 with the same results in the case of median values. Kruskal-Wallis test showed there is statistically significant difference between the samples (p = 0,0003) and there are number of pairs of samples that are differing significantly from one another: (Z1, Z2), (Z1, Z3), (Z1, Z5), (Z1, Z6), (Z2, Z5), (Z3, Z5), (Z4, Z5), (Z5, Z6), (Z5, Z7), (Z5, Z8), (Z5, Z9), (Z5, Z10) and (Z6, Z9).

Table 2 Reading speed and subjective difficulty of reading test results

G/R	T1	Z1	T2	Z2	T3	Z3	T4	Z4	T5	Z5	T6	Z6	T7	Z7	T8	Z8	T9	Z9	T10	Z10	T11	Z11
1	30,99	6	37,01	10	31,09	7	33,19	9	42,85	5	30,53	8	39,99	6	28,65	7	40,15	9	40,37	8	23,93	10
2	18,19	8	40,72	9	47,90	7	27,93	6	30,47	5	36,73	6	37,23	6	39,29	5	27,90	6	38,74	7	24,86	7
3	42,50	7	25,91	8	32,68	6	28,56	7	32,74	5	29,51	7	32,16	7	35,50	7	36,46	7	33,90	6	36,93	5
4	30,55	8	51,64	8	32,19	8	24,07	9	43,03	7	32,35	8	35,48	4	37,93	8	37,88	9	28,82	10	27,07	7
5	16,13	7	38,61	7	32,71	8	28,87	9	26,86	9	70,03	2	43,07	5	28,17	8	30,23	3	24,33	7	38,47	6
6	38,34	8	43,06	9	30,50	7	28,08	6	21,57	6	40,89	8	30,77	8	29,50	10	33,32	8	36,11	8	35,41	7
7	25,74	10	40,66	6	42,73	7	37,67	7	47,63	9	29,22	9	39,44	9	33,60	10	32,82	9	28,27	7	39,24	7
8	23,00	9	22,93	9	40,82	9	20,51	9	56,78	5	43,35	3	39,70	9	27,08	6	26,11	3	36,88	7	39,25	8
9	27,31	10	38,61	9	38,34	8	26,93	10	41,29	5	30,37	8	26,28	8	42,55	6	40,58	9	30,66	8	43,28	5
10	29,64	9	43,50	7	19,41	6	31,42	8	29,67	8	31,23	8	30,90	10	34,88	5	29,52	10	24,14	7	33,83	7
11	35,41	7	36,86	6	31,75	4	24,93	9	46,73	7	25,22	7	23,92	8	28,20	6	24,25	9	27,68	9	27,69	9
12	31,35	9	40,13	5	24,92	6	31,04	6	63,55	6	42,75	7	31,01	8	32,01	9	27,44	8	23,51	10	32,37	6
13	35,77	7	39,60	5	25,67	5	33,80	8	32,85	4	27,17	8	34,22	9	47,15	8	34,99	10	40,03	8	24,64	7
14	33,14	8	28,21	7	24,25	10	42,36	6	40,58	4	33,93	8	30,67	10	32,13	9	22,33	10	34,18	8	25,89	8
15	21,19	9	24,60	8	46,81	9	29,01	9	35,07	3	28,61	9	34,57	6	45,92	7	27,60	10	36,68	8	19,95	7
16	32,01	8	29,41	5	16,71	6	32,82	8	37,51	5	27,74	7	28,06	7	29,30	8	25,10	8	25,01	10	39,33	7
17	24,34	9	35,56	7	35,31	8	33,94	6	35,59	5	29,73	5	30,38	9	24,58	9	30,40	6	29,26	7	32,50	6
18	27,89	10	47,27	6	37,25	5	38,31	6	42,69	5	29,66	7	41,80	7	31,39	7	37,91	7	31,22	7	30,90	6
19	33,31	8	40,18	4	22,76	8	28,13	8	37,01	6	29,16	7	33,40	7	42,01	5	29,12	8	32,53	7	26,84	8
20	29,27	8	26,40	7	42,48	6	31,62	7	35,01	6	31,66	7	38,69	6	26,93	10	30,42	8	21,66	9	39,50	7

Following the data analysis, it would be safe to conclude that the best results considering the speed of reading and subjective difficulty of reading, across the various parameters, have been achieved in the Group 1 that was reading single column 12 pt Minion Pro text spread printed on paper. The lowest results across the parameters have been found for the results in Group 5 that

was reading two columns 12 pt Minion Pro text spread displayed on iPad.

Table 3 Results of the statistical tests for reading speed (T) and subjective difficulty of reading (Z) for samples 1 to 10

Sample	Mean (s)	Median (s)	Kruskal-Wallis test	
			H	*p*
T1	29,3	30,09	23,2178	0,0057
T2	36,54	38,61		
T3	32,81	32,43		
T4	30,66	30,02		
T5	38,97	37,26		
T6	33,99	30,45		
T7	34,09	33,81		
T8	33,83	32,07		
T9	31,23	30,32		
T10	31,2	30,94		
Sample	Mean (s)	Median (s)	Kruskal-Wallis test	
			H	*p*
Z1	8,25	8	30,5793	0,0003
Z2	7,1	7		
Z3	7	7		
Z4	7,65	8		
Z5	5,75	5		
Z6	6,95	7		
Z7	7,45	7,5		
Z8	7,5	7,5		
Z9	7,85	8		
Z10	7,9	8		

Compared between the three (T1, Z1;T2, Z2; T3, Z3), the results showed reading on paper (T1, Z1) is significantly faster and easier in comparison to either iPad (T3, Z3) or computer screen (T2, Z2) with the later significantly slower and difficult than both paper and iPad. It is important to note that there is no significant difference between samples T1 and T3 with regard to reading speed. There is, however, significant difference between Z1 and Z2 and Z3, with Z3 sample achieving the lowest mean value within the group, but without any statistically significant difference between Z2 and Z3.

When comparing median values, four samples are standing up with the lowest results for reading speed and highest results for subjective difficulty of reading –the samples 1, 4, 9 and 10, where T4 is the lowest median value for reading speed and all four have a median value of 8 for reading difficulty (Z), although, mean value Z1 is the only result above 8 and Z10 is closest to 8 in comparison to all the other groups. All four samples' medians for reading speed are below 31 s.

5 Conclusion

When reading a text on iPad, several findings presented in this article seem important as a guideline for designing a text.

First, sans serif Gotham font is more readable and legible than Minion Pro serif font. Although, when presented in bigger font sizes (e.g. 14 and 16 px) there is no significant difference between the two letter cases. However, without further research with different serif and sans serif fonts it is hard to make any definite conclusions. If this difference holds out, conclusion would be that what was advantage in one display technology – serifs in printing – seems not to be transferred in another display technology, in this case into iPad. When median values are examined, it is obvious that texts presented in Gotham sans serif font are both in reading speed and reading difficulty quite close to quality of reading from paper. The only exemption, but only when it comes to reading speed, is the text presented in 10 px Minion Pro one column spread on iPad, results for which are comparable to Gotham spreads. However, reading difficulty for this piece was second lowest across all pieces.

Second, two column spread of text was the hardest to read, probably because the number of characters in a row was too low for continuous reading, as suggested in previous research [see 3,4 and 5]. Since it was not in focus of this paper, further research would be needed to determine the character per row threshold when preparing longer texts intended for reading on iPad display. From the scope of this experiment, layout of the text with 79 characters per row and above should be readable and legible enough for reading texts on iPad.

Third, if the findings presented in this paper, despite some obvious research limitations – e.g. sample size, sample demographics and layout limitations – are confirmed by further research, it would be obvious that iPad display technology would be able to replace paper for reading longer texts in series of formats, something PC screen is still unable to achieve. Trends in newspaper, magazines and books publishing [24] are to some extent suggesting precisely that conclusion. For this technological substitution, in comparison to technical characteristics of the reading gadgets, long term habits of consumers seem to be more important, as some research is already suggesting [25, 26].

6 References

[1] Gelb, I. J. A study of writing, University of Chicago Press, Chicago, 1963

[2] Daniels, P. T.; Bright, W. (eds.) The World's Writing Systems', Oxford University Press, Oxford, New York, 1996

[3] Tinker, M. Legibility of print, Iowa state University Press, Ames, IA, 1963

[4] Rabinowitz, T. Exploring Typography, Thomson/Delmar Learning, Clifton Park, NY, 2006

[5] McMurtrie, D. The Philosophy of Modernism in Typography. // in Looking Close 3: Critical Writings on Graphic Design, Skyhorse Publishing, New York, 1929/1999.

[6] Tarasov, D. A. et al. Legibility of textbooks: a literature review. // Procedia - Social and Behavioral Sciences. 174 (2015), pp. 1300-1308. https://doi.org/10.1016/j.sbspro.2015.01.751

[7] Dillon, A. Reading from paper versus screens: a critical review of the empirical literature. // Ergonomics. 35, 10(1992), pp. 1297-1326. https://doi.org/10.1080/00140139208967394

[8] Askwall, S. Computer supported reading vs reading text on paper: a comparison of two reading situations. // International Journal of Man-Machine Studies. 22, (1985), pp. 425-439. https://doi.org/10.1016/S0020-7373(85)80048-1

[9] Oborne, D.; Holton, D. Reading from screen versus paper: there is no difference. // International Journal of Man-Machine Studies. 28, 1(1988), pp. 1-9. https://doi.org/10.1016/S0020-7373(88)80049-X

[10] Noyes, J. M.; Garland, K. J. Computer vs. paper-based tasks: Are they equivalent? // Ergonomics. 51, 9(2008), pp. 1352-1375. https://doi.org/10.1080/00140130802170387
[11] Weisenmiller, E. M. A study of the readability of on-screen text. // Doctoral dissertation, Virginia Polytechnic Institute and State University, July, 1999
[12] Baker, R. D. Comparing the readability of text displays on paper, e-book readers, and small screen devices. // Doctoral dissertation, University of North Texas, May 2010
[13] Ali, Ahmad Zamzuri Mohamad et al. Reading On The Computer Screen: Does Font Type has Effects on Web Text Readability? // IES 6.3 (2013): Retrieved from Web. 11. Jan. 2016
[14] Hughes, L. E.; Wilkins, A. J. Reading at a distance: Implications for the design of text in children's big books. // British Journal of Educational Psychology. 72, (2002), pp. 213-226. https://doi.org/10.1348/000709902158856
[15] Sheedy, J. E.; Subbaram, M. V.; Zimmerman, A. B.; Hayes J. R. Text legibility and letter superiority effect. // Human Factors: The Journal of the Human Factors and Ergonomics Society. 47, 4(2005), pp. 797-815.
[16] Chandler, S. B. Running Head: Legibility and comprehension of onscreen type. // Doctoral dissertation, Virginia Polytechnic Institute and State University, 2001.
[17] Subbaram, V. M. Effect of display and text parameters on reading performance. // Doctoral dissertation, Ohio State University, 2004
[18] Just, M. A.; Carpenter, P. A. A theory of reading: From eye fixations to comprehension. // Psychological Review. 87, (1980), pp. 329-355. https://doi.org/10.1037/0033-295X.87.4.329
[19] Gasser, M.; Boeke, J.; Haffernan, M.; Tan, R. The Influence of Font Type on Information Recall. // North American Journal of Psychology. 7, 2(2005), pp. 181-188.
[20] Nielsen, J. iPad and Kindle reading speeds, 2010. http://www.useit.com/alertbox/ipad-kindle-reading.html (11.1.2016)
[21] Brangan, S. Razvoj formula citkosti za zdravstvenu komunikaciju na hrvatskom jeziku (Developing readability formulas for healthcare communication in Croatian language). // Doctoral dissertation, Medicinski fakultet Sveučilišta u Zagrebu, 2011
[22] Čerepinko, D.; Mrvac, N.; Milković, M. Očekivanja i zahtjevi potencijalnih korisnika grafičkog korisničkog sučelja "tablet novina"' (User expectations and requirements for "tablet newspaper" oriented graphic interfaces). // Tehnički glasnik. 9, 1(2015), pp. 84-87.
[23] McGill, R.; Tukey, J.; Larsen, W. Variations of Box Plots. // The American Statistician. 32, 1(1978), pp. 12-16. https://doi.org/10.1080/00031305.1978.10479236
[24] The Magazine Media Factbook 2015. MPA – The Association of Magazine Media, New York, 2015.
[25] Garcia, M. iPad Design Lab - Basic: Storytelling in the Age of the Tablet, Blue Ash, OH: HOW Books, 2012.
[26] Hollander, B. A.; Krugman, D. M.; Reichert, T.; Avant, J. A. The e-reader as replacement for the print newspaper. // Publishing Research Quarterly. 27, (2011), pp. 126-134. https://doi.org/10.1007/s12109-011-9205-8

Authors' addresses

doc. dr. sc. Darijo Čerepinko, dipl. ing.
v. pred. Damira Keček, dipl. ing.
v. pred. Mario Periša, dipl. ing.
University North
104. brigade 3, 42000 Varaždin, Croatia
darijo.cerepinko@unin.hr
damira.kecek@unin.hr
mario.perisa@unin.hr