
Evaluation of Efficient Line Lengths for Better Readability

ZAHID HUSSAIN*, SYED ASIF ALI SHAH**, AND WAJIHA SHAH***

RECEIVED ON 30.09.2011 ACCEPTED ON 01.12.2011

ABSTRACT

In this paper the major findings of a formal experiment about onscreen text line lengths are presented. The experiment examined the effects of four different line lengths on the reading speed and the reading efficiency. Efficiency is defined as a combination of reading speed and accuracy. Sixteen people between the age of 24 and 36 participated at the experiment. The subjects had to read four different texts with an average line length around 2000 characters. The texts contained substitution words, which had to be detected by the subjects to measure reading accuracy. Besides objective measures like reading speed and accuracy, the subjects were asked to subjectively vote on their reading experience. The results from our objective measures show strong similarities to those of the work done previously by different researchers. The absolute reading speed grows when the line length grows from CPL (Characters Per Line) 30-120. The measured reading efficiency, however, doesn't grow steadily, although a growing trend can be seen. This is due to the fact, that the test persons found in average more substitution words from the 60 CPL text than they did from the 30 and 90 CPL texts. The reading speed seems to increase while the line length increases but the overall comprehension seems to peak at medium line lengths. As in the previous studies, our test persons also prefer the medium (60 and 90 CPL) line lengths, although they perform better when reading longer lines. In the overall subjective opinion 13 out of 16 test persons selected the 60 or 90 CPL line length as their favorite. The literature doesn't truly provide a scientific explanation for the difference between the objective performance and the subjective preference. A natural hypothesis would be that the line length that is the fastest to read would also feel most comfortable to the readers but in the light of this and the earlier research it seems like this is not the case.

Key Words: Line Length, Reading Efficiency, Usability Testing.

1. INTRODUCTION

Researchers have been interested in finding the optimal line length since end of the 19th century. The research was concentrated on book and newspaper readability until in the 1980's the focus began to shift towards on-screen readability due to growing demand. The optimal line length is certainly different for newspapers, books and online information. Newspapers are most often using narrow line lengths. More wide lines are used in books and the on-screen line lengths vary a lot

due to different sized monitors. The first of the studies by Weber, [1]. Weber, A., came into a conclusion that the optimal line length for newspapers and books is four inches and in no case should it exceed six inches. Just about 50 years later, Tinker, et. al. [2] claimed that the optimal line length was between 3 and 3.5 inches. They also suggested that if line length is longer than 7.3 inches, the reader might not find the next line after finishing the previous one.

* Assistant Professor, Department of Information Technology, Quaid-e-Awam University of Engineering, Science & Technology, Nawabshah.

** Assistant Professor, Department of Electrical Engineering, Mehran University of Engineering & Technology, Jamshoro.

*** Assistant Professor, Department of Electronic Engineering, Mehran University of Engineering & Technology, Jamshoro

The first ones to investigate the impact of line length to on-screen readability were Ducknick, et. al. [3] whose results showed that text scaled to full screen width was significantly faster to read than text that filled only a third of the screen width [3]. In a study by Dyson, et. al. [4] suggested that the reading rate increases, with increasing number of letters per line. The study used a 12 point font. The slowest reading speed was noticed when line length was 4 inches, the fastest at 7.3 inches. But their test-readers preferred line lengths of 4 inches. Bernard, M., et. al., tested three different line lengths (3.3, 5.7 and 9.6 inches) with a 12 point sized font on a prose-text in [5]. Their results show no reliable differences in average reading speed between these three lengths. Shaikh, [6] also conducted the line length effect study for reading online news. Recently Holzinger, et. al. [7] has carried out a study for paper vs. screen with medical practitioners concluding that paper still has importance in people's lives.

In short one could say that, concerning the factual results, there is a tendency that a longer line length correlates with faster reading speed and that a medium line length correlates with a better overall comprehension. In most studies, however, the differences that were found are often non-significant. There is an agreement that lines shorter than 2.6 inches are not reasonable. Subjectively the users tend to prefer a medium line length (about 60 CPL), Aschbacher, et. al. [8]. The aim of this study was to further investigate the effect of line length to on-screen readability.

2. TEST METHODOLOGY

2.1 User Profiles

Because the study discussed on-screen readability and particularly reading from the browser window the targeted potential user group is all Internet users. For practical reasons the subjects' age was limited between 18 and 40 years. The subjects were also expected to have some experience in computer use, to have normal or corrected-to-normal vision and to be native German speakers, as shown in Table 1.

Sixteen subjects volunteered for this study (13 Men and 3 Women). They ranged in age from 24-36 (Median 27.5) years. The average computer usage experience was 16.0 years (Median 15). The average computer work per week was 51.43 (Median 50). Most of the participants were employees or students; so they have lot of experience of on-screen reading. Most of them were working with MS

Windows and Firefox and most of them had already participated in other usability tests. Most of them preferred to magnify the text for reading online. The majority read newspapers online.

2.2 Test Design

The study used 4x4 factorial designs with line length and text as independent variables. The experiment was balanced through all of the 16 test users with a fully balanced Latin square design.

To determine the suitable line lengths for the study different commonly used web news pages were examined. Based on this examination following line lengths were selected: 30, 60, 90, and 120 CPL. The four texts were selected from www.derStandard.at (An Austrian Newspaper). Each of the texts has a length of approximately 2000 characters, as shown in Table 2.

In each text 10 words are replaced with other similar words. These substitution words are used to measure reading accuracy. The subjects had to find these words while reading the text samples. An example of a substitution word is rest->test, where a word is substituted with another, similar-sounding word, which doesn't make any sense in the context. Measured values were reading time, reading speed and reading efficiency. The efficiency is calculated as:

$$\text{Efficiency} = (\text{Reading Speed}) * (\% \text{ Found Substitution Words})$$

Reading speed is measured in words per minute. Four different line lengths and text samples result in a 4x4 matrix with a total number of 16 different test cases. These test cases are arranged in a balanced Latin square as shown in Table 3.

2.3 Test Tasks

In this section we describe the test process. Each subject was assigned to a random test case, by picking a number from a bag. After that they had to fulfill an example test case, which introduced the test scenario. If the subject had no more questions about the test process, she/he was asked to pick the assigned test case from a dropdown menu. The first text was displayed and the subject had to read the text and speak out loud the substitution words. As the subject finished reading she/he had to push the "Fertig (Finish)" button. After that she/he could continue to the next text, by pushing the "Start" button. The time

between the start of a text and the end was measured by the server side protocol of the test environment. On the software side the test environment was set up with the help of a XAMPP Server. The texts were displayed as HTML with the use of the Firefox Web browser maximized to the computer's native resolution. The texts were rendered with the font Verdana in size 12 point. The text color was black (#000000) on a slightly gray background (#F0F0F0). The dedicated line lengths were created by tables with fixed widths and the Firefox native line break. The texts and all other necessary data was stored in a database by the PHP server, which stored the measured timing either. The texts were taken from the online news portal derStandard.at. All 4 texts are tourist reports from different locations. They are all around 2000 characters. The reference from the Latin square distribution is mentioned in the brackets. The introduction text was presented to each subject before the test.

2.4 Test Environment

The tests were recorded with a digital video camera and an external microphone. The test utility software was installed on an Acer Aspire 1710 Laptop, which has an Intel P4 (3.4 GHz) processor, 1 GB RAM, a 17" LCD-display with resolution of 1280x1024 and 16.7 million colors. The

operating system was Windows XP Professional SP2. During the actual tests there were three facilitators in the room in addition to the actual test person: one was operating the camera, one the microphone and one was running the tests with the test person.

3. RESULTS

The measurement was divided into two parts. The first part is the objective measurement, which is a combination of the reading speed and the correct substitution words found. The second part is the subjective measurement, which reflects the subjects own experience. The objective measurement performed with a steady increase in reading speed and also an increase in efficiency. The correct found substitution words had an average around 6.5 out of 10 for all 4 line lengths.

TABLE 2. TEXTS CONTAINING CHARACTERS AND WORDS

Taxt No.	Characters	Words
1	2070	280
2	1985	290
3	1950	204
4	1997	264

TABLE 1. PARTICIPANTS' DESCRIPTION

No.	Gender	Age	Profession	Vision	Color Blind	Education	PC Exp (Years)	PC Work/Week	OS	Browser	Usability Tests Experience
TP01	M	27	Employee	Glasses	No	University	13	40	Windows	Firefox	No
TP02	F	28	Employee	Contact lens	No	University	16	55	Windows	Firefox	Yes
TP03	M	36	Employee	Glasses	No	University	20	30-60	Windows	Firefox	Yes
TP04	M	30	Employee	N/A	No	University	18	60	Unix/Linux	Firefox	Yes
TP05	M	24	Employee	N/A	No	High School	14	40	Windows	Firefox	Yes
TP06	M	29	Employee	N/A	No	University	15	60	Unix/Linux	Firefox	No
TP07	M	24	Student	N/A	No	High School	11	50	Windows	Firefox	Yes
TP08	M	26	PhD Student	N/A	No	University	15	80	Windows	Firefox	Yes
TP09	M	28	Employee	N/A	No	University	15	40	Unix/Linux	Firefox	Yes
TP10	M	30	Employee	N/A	Red, Green	University	20	50	Windows	Firefox	No
TP11	F	25	Psychologist	N/A	No	University	10	50	Windows	Firefox	Yes
TP12	M	28	Student	Glasses	No	University	20	30-40	MacOS X	Firefox	No
TP13	F	25	Web Developer	N/A	No	College	15	50	Windows	Firefox	Yes
TP14	M	26	Student	N/A	No	High School	20	25	Windows	Opera	Yes
TP15	M	34	Employee	Glasses	No	Doctorate	20	40	Windows	Firefox	Yes
TP16	M	24	Student	N/A	No	High School	14	60	MacOS X, Unix/Linux	Firefox, Safari	Yes

The subjective measures showed, that most people liked a line length between 60 and 90 because 43.7% preferred the 60 CPL and 37.5% liked the 90 CPL the most, which is more than 80% of the test subjects. These figures correlate with the subjective answers from reading speed, reading accuracy and readability.

3.1 Objective Measures

In this section we will analyze our measurements of reading speed and substitution words found. The Tables 4-7 represent the 4 different states of the measured variable of interest. For each value, we gathered the time it took a subject to read the text and the substitution words found. The timing was automatically done by the computer. The substitution words were written down, by the moderator of the test. False recognized substitution words were not taken into account, as they are not interested for our study. (The substitution words are present to ensure a valuable reading; in addition to that false recognized words are still an indicator for valuable reading). The gathered values were statistically analyzed. The most important calculation in the Tables 4-7 is efficiency. This value is the overall time it took a subject to read the text according to the length of

the text and the substitution words found. With this value, we can say, how fast and accurate a subject read the text. Moreover, all other variable values, like text length or substitution words are equally distributed. An analysis over the actual found values is described after the tables. Formula for the efficiency algorithm is given by:

$$\text{Efficiency} = \frac{\text{Found Substitution Words} * \left(\frac{\text{Words}}{\text{Time}} \right)}{\text{Total Substitution Words}}$$

Overall it can be said, that the reading speed is going up for longer line lengths, but not steadily. The difference of around fifteen seconds from 60 CPL (Table 5) to 120 CPL (Table 7) is quite remarkable as this is a boost of around 9%. Moreover we found out, that the accuracy in found substitution words has its peak at the longest line either (Table 7).

After taking into account the overall length of the text and the absolute reading speed, the efficiency showed a steady rising from 30-120 CPL. So the expansion of the line length from 30-120 CPL brought an increase in efficiency of approximately 13% (Tables 4 and 7). The standard deviation of the efficiency has a minimum in the 120 CPL test with 17.94 (Table 7) and a maximum at 60 CPL with 29.54 (Table 5), which is nearly 40% of the mean value.

3.2 Subjective Measures

Following feedback survey Tables 8-11 show readability, reading speed, reading accuracy and preferred line length.

The results from the feedback questionnaires show that our test participants preferred 60 CPL (7 out of 16 participants), followed by 90 CPL (6 out of 16 participants). The worst preferred text was 30 CPL (1 out of 16 participants); Table 11. For the good readability of texts, 60 CPL text was rated with an average of 3.94 (Table 8), followed by 90 CPL (an average of 3.50). The 120 CPL text had the worst readability rating (an average of 2.81). Again 60 CPL text was rated with an average of 3.81 as having good reading speed, followed by 90 CPL (an average of 3.38); Table 9. The 120 CPL text again had the worst reading speed rating (an average of 2.81). But 90 CPL text was rated with an average of 3.75 as having good reading accuracy, followed by 60 CPL (an average of 3.38). 30 CPL text had the worst reading accuracy (an average of 2.25), see Table 10.

TABLE 3. BALANCED LATIN SQUARE

Test No.	T1	T2	T3	T4
1	60	30	90	120
2	30	120	60	90
3	120	90	30	60
4	90	60	120	30
Test No.	T2	T1	T4	T3
5	60	30	90	120
6	30	120	60	90
7	120	90	30	60
8	90	60	120	30
Test No.	T3	T4	T1	T2
9	60	30	90	120
10	30	120	60	90
11	120	90	30	60
12	90	60	120	30
Test No.	T4	T3	T2	T1
13	60	30	90	120
14	30	120	60	90
15	120	90	30	60
16	90	60	120	30

TABLE 4. 30 CHARACTERS PER LINE

User No.	Test Case No.	Time (Seconds)	Words Found	Text	Characters	Words	Charracters/ Minimum	Words/ Minimum	Efficiency
4	1	193	6	T2	1985	290	617	90	54
6	2	120	8	T1	2070	280	1035	140	112
7	3	181	8	T3	1950	294	646	97	78
10	4	148	6	T4	1997	264	810	107	64
8	5	126	4	T1	2070	280	986	133	53
15	6	220	8	T2	1985	290	541	79	63
2	7	102	7	T4	1997	264	1175	155	109
9	8	145	5	T3	1950	294	807	122	61
3	9	175	6	T4	1997	264	685	91	54
13	10	238	5	T3	1950	294	492	74	37
1	11	175	8	T1	2070	280	710	96	77
16	12	234	7	T2	1985	290	509	74	52
5	13	158	9	T3	1950	294	741	112	100
12	14	106	4	T4	1997	264	1130	149	60
14	15	134	5	T2	1985	290	889	130	65
11	16	106	6	T1	2070	280	1172	158	95
Mean		160.06	6.38				808.95	113.03	70.93
Standard Deviation		44.83	1.54				232.80	28.79	22.20

TABLE 5. 60 CHARACTERS PER LINE

User No.	Test Case No.	Time (Seconds)	Words Found	Text	Characters	Words	Charracters/ Minimum	Words/ Minimum	Efficiency
4	1	193	9	T1	2070	280	644	87	78
6	2	116	9	T3	1950	294	1009	152	137
7	3	179	7	T4	1997	264	669	88	62
10	4	189	6	T2	1985	290	630	92	55
8	5	175	6	T2	1985	290	681	99	60
15	6	187	6	T4	1997	264	641	85	51
2	7	88	6	T3	1950	294	1330	200	120
9	8	152	8	T1	2070	280	817	111	88
3	9	191	5	T3	1950	294	613	92	46
13	10	229	7	T1	2070	280	542	73	51
1	11	190	7	T2	1985	290	627	92	64
16	12	193	4	T4	1997	264	621	82	33
5	13	164	6	T4	1997	264	731	97	58
12	14	123	8	T2	1985	290	968	141	113
14	15	128	7	T1	2070	280	970	131	92
11	16	95	5	T3	1950	294	1232	186	93
Mean		162.00	6.63				795.19	113.07	75.12
Standard Deviation		40.59	1.41				237.47	38.29	29.54

TABLE 6. 90 CHARACTERS PER LINE

User No.	Test Case No.	Time (Seconds)	Words Found	Text	Characters	Words	Charracters/ Minimum	Words/ Minimum	Efficiency
4	1	158	8	T3	1950	294	741	112	89
6	2	101	5	T4	1997	264	1186	157	78
7	3	185	6	T2	1985	290	644	94	56
10	4	187	9	T1	2070	280	664	90	81
8	5	109	3	T4	1997	264	1099	145	44
15	6	160	7	T3	1950	294	731	110	77
2	7	91	7	T1	2070	280	1365	185	129
9	8	127	3	T2	1985	290	938	137	41
3	9	201	5	T1	2070	280	618	84	42
13	10	205	8	T2	1985	290	581	85	68
1	11	198	9	T4	1997	264	605	80	72
16	12	226	6	T3	1950	294	518	78	47
5	13	136	5	T2	1985	290	876	128	64
12	14	123	9	T1	2070	280	1010	137	123
14	15	126	6	T3	1950	294	929	140	84
11	16	113	4	T4	1997	264	1060	140	56
Mean		152.88	6.25				847.76	118.80	71.98
Standard Deviation		42.66	2.02				248.23	31.89	26.26

TABLE 7. 120 CHARACTERS PER LINE

User No.	Test Case No.	Time (Seconds)	Words Found	Text	Characters	Words	Charracters/ Minimum	Words/ Minimum	Efficiency
4	1	172	9	T4	1997	264	697	92	83
6	2	123	6	T2	1985	290	968	141	85
7	3	163	9	T1	2070	280	762	103	93
10	4	151	7	T3	1950	294	775	117	82
8	5	98	3	T3	1950	294	1194	180	54
15	6	189	8	T1	2070	280	657	89	71
2	7	96	6	T2	1985	290	1241	181	109
9	8	120	5	T4	1997	264	999	132	66
3	9	189	6	T2	1985	290	630	92	55
13	10	208	9	T4	1997	264	576	76	69
1	11	158	9	T3	1950	294	741	112	100
16	12	255	8	T1	2070	280	487	66	53
5	13	105	6	T1	2070	280	1183	160	96
12	14	100	6	T3	1950	294	1170	176	106
14	15	150	8	T4	1997	264	799	106	84
11	16	86	4	T2	1985	290	1385	202	81
Mean		147.69	6.81				891.39	126.60	80.40
Standard Deviation		47.56	1.87				273.57	42.24	17.94

TABLE 8. READABILITY (0-6 SCALE)

No.	30 CPL	60 CPL	90 CPL	120 CPL
TP01	4	6	2	0
TP02	4	5	0	2
TP03	6	5	3	2
TP04	1	3	3	2
TP05	1	2	6	4
TP06	0	1	5	5
TP07	6	5	2	1
TP08	4	6	4	3
TP09	2	3	2	3
TP10	5	4	2	1
TP11	1	3	5	3
TP12	3	4	6	5
TP13	4	3	4	3
TP14	1	6	5	4
TP15	1	2	5	4
TP16	3	5	2	3
Average	2.88	3.94	3.50	2.81
Standard Deviation	1.93	1.57	1.75	1.42

TABLE 9. READING SPEED (0-6 SCALE)

No.	30 CPL	60 CPL	90 CPL	120 CPL
TP01	5	5	2	1
TP02	4	6	2	5
TP03	4	4	3	2
TP04	1	3	4	2
TP05	2	3	5	4
TP06	0	0	6	6
TP07	5	4	1	0
TP08	5	6	4	2
TP09	2	3	2	3
TP10	5	4	2	1
TP11	1	3	5	3
TP12	3	4	4	3
TP13	4	4	3	2
TP14	0	6	5	4
TP15	1	1	5	4
TP16	4	5	1	3
Average	2.88	3.81	3.38	2.81
Standard Deviation	1.86	1.68	1.59	1.56

TABLE 10. READING ACCURACY (0-6 SCALE)

No.	30 CPL	60 CPL	90 CPL	120 CPL
TP01	3	3	3	3
TP02	2	6	1	5
TP03	4	4	3	1
TP04	1	4	5	3
TP05	0	2	4	3
TP06	0	0	6	6
TP07	4	5	2	1
TP08	2	6	5	3
TP09	3	2	3	2
TP10	5	4	3	2
TP11	0	1	5	3
TP12	3	4	5	5
TP13	2	2	3	4
TP14	0	5	6	4
TP15	2	2	5	4
TP16	5	4	1	4
Average	2.25	3.38	3.75	3.31
Standard Deviation	1.73	1.75	1.61	1.40

TABLE 11. PREFERRED LINE LENGTH

No.	30 CPL	60 CPL	90 CPL	120 CPL
TP01		1		
TP02		1		
TP03		1		
TP04			1	
TP05			1	
TP06				1
TP07	1			
TP08		1		
TP09			1	
TP10		1		
TP11			1	
TP12			1	
TP13		1		
TP14			1	
TP15				1
TP16		1		
Percentage	6.25%	43.75%	37.50%	12.50%

4. CONCLUSIONS

The results from our objective measures show strong similarities to those of the work done previously by different researchers. The absolute reading speed grows when the line length grows from 30-120 CPL. The measured reading efficiency, however, doesn't grow steadily, although a growing trend can be seen. This is due to the fact, that the test persons found in average more substitution words from the 60 CPL text than they did from the 30 and 90 CPL texts. This indicates a similar conclusion: The reading speed seems to increase while the line length increases but the overall comprehension seems to peak at medium line lengths. As in the previous studies, our test persons also prefer the medium (60 and 90 CPL) line lengths, although they perform better when reading longer lines. In the overall subjective opinion 13 out of 16 test persons selected the 60 or 90 CPL line length as their favorite. The literature doesn't truly provide a scientific explanation for the difference between the objective performance and the subjective preference. A natural hypothesis would be that the line length that is the fastest to read would also feel most comfortable to the readers but in the light of this and the earlier research it seems like this is not the case. As a difference is found between the objective and subjective results, the final decision has to be made by the web designers: do they give more value to the actual reading speed or the subjective preference of the readers. The difference in reading efficiency between 60 or 90 CPL texts and the 120 CPL text was not drastic but only 2 test persons out of 16 selected the 120 CPL line length as their preference. Because the web is among other things a highly visual media it could be suggested that the benefits gained with the longer line length are less important than the losses it causes in overall user preference. Although the differences found in our study are not obvious, one could suggest based on our results that for general web use the medium line lengths (close to 60-90 CPL) are a safe compromise between subjective user preference and overall reading efficiency.

Future work will concentrate on conducting the same type of study in the context of Pakistani users for reading newspapers like Dawn, etc. as well as reading on the tiny screens of hand-held mobile devices and tablets.

ACKNOWLEDGEMENTS

The authors are highly thankful to Higher Education Commission of Pakistan, Technical University of Graz, Austria, Vienna University of Technology, Austria, Quaid-e-Awam University of Engineering, Science & Technology, Nawabshah, and Mehran University of Engineering & Technology, Jamshoro, Pakistan, for providing necessary funding and research facilities during their research studies.

REFERENCES

- [1] Weber, A., "Ueber Die Augenuntersuchungen in Den Hoheren Schulen zu Darmstadt", Abtheilung fur Gesundheitspflege, [Cited in Tinker and Paterson, 1929], March, 1881.
- [2] Tinker, M.A., and Paterson, D.G., "Studies of Typographical Factors Influencing Speed of Reading: Length of Line", The Journal of Applied Psychology, Volume 13, No. 3, pp. 205-219, 1929.
- [3] Duchnicky, J.L., and Kolars, P.A., "Readability of Text Scrolled on Visual Display Terminals as a Function of Window Size", Human Factors Journal, Volume 25, No. 6, pp. 683-692, 1983.
- [4] Dyson, M.C., and Kipping, G.J., "The Effects of Line Length and Method of Movement on Patterns of Reading from Screen", Visible Language Journal, Volume 32, No. 2, pp. 150-181, 1998.
- [5] Bernard, M., Fernandez, M., and hull, S., "The Effects of Line Length on Children and Adults' Online Reading Performance", Usability News Web Newsletter, Volume 4, No. 2, 2002.
- [6] Shaikh, A.D., "The Effects of Line Length on Reading Online News", Usability News Web Newsletter, Volume 7, No. 2, 2005.
- [7] Holzinger, et. al., "Investigating Paper vs. Screen in Real-Life Hospital Workflows: Performance Contradicts Perceived Superiority of Paper in the User Experience", International Journal of Human-Computer Studies, Volume 69, No. 9, August, 2011.
- [8] Aschbacher, H., Fressel, C., and Gritschacher, T., and Zimmermann, K., "Line Length", 706.041 Information Architecture and Web Usability WS 2007/2008, IICM, Technical University Graz, Austria, 2008.