原　著

# How Well the Elderly Evaluate the Readability of E-paper Devices: The Effects of Font Sizes

# 高齢者は電子ペーパー機器をどのくらい読みやすいと評価しているか？ ---- 文字サイズの効果について

Paul Lege[1], Shigusa Matsunami[1], Takehito Kojima[2], Masaru Miyao[1]
1) Nagoya University, 2) Chubu-Gakuin University

## Abstract

The purpose of this experiment was to measure and understand reading tests how the elderly evaluated the effects of the three font sizes using the Courier typeface from two different displays. The two devices included the E-paper device (Amazon Kindle Deluxe™) and a backlit LCD (Apple iPad™). These were compared to conventional paper as a reference. The study involved 99 participants of various ages, who were given a short reading task that measured for speed and a correct answer rate from different sets of font sizes of 8 pt (2.75 mm), 12 pt (3.25 mm), and 16 pt (5.75 mm). After each task, the readers rated the readability of visual performance of the device at each of these point sizes. Our study found that the readability of the E-paper was equal to the performance of paper text. Therefore, E-paper designed to function similar to paper can contribute to improving readability including magnification function, especially for the elderly. On the basis of this study, we recommend a default setting higher than 8 pt (2.75 mm) be displayed on the screens of E-papers.

## 抄録

この実験の目的は、電子ペーパーに表示された文章を読む課題について、高齢者を含む多数の実験参加者によって、どのくらい電子ペーパー機器の読みやすさを評価しているかについて検証したものである。2 種の情報機器（電子ペーパーのアマゾン・キンドル・デラックスとバックライト型液晶のアップル・アイパッド）を使用し、クーリエ・フォントを使った 3 つのフォントサイズの効果を評価した。対照として従来の紙も比較にために用いた。実験には、幅広い年齢層を含む 99 人の実験参加者が調査された。彼らは、短い読み課題を与えられた。このテストは 8pt（2.75mm）、12pt（3.25mm）、および 16pt（5.75mm）のフォントサイズの文字を読ませ、スピードと正答率を測定した。個々の機器や紙での読み課題の後に、参加者は、機器や紙のそれぞれ 3 段階の文字サイズごとに、読みやすさを評価した。その結果、情報機器の読みやすさは、紙と同等であることがわかった。

したがって、紙とよく似た性質を有する電子ペーパーは、文字も拡大も容易で、高齢者にとって、読みやすさを改善することに寄与できる。この研究に基づいて、私達は、電子パーパーの画面において、8pt（2.75mm）の文字高を最低基準とする読みやすさに関する設定を推奨する。

Keywords: Readability of E-paper, Font Sizes, Aging, Standards
キーワード：電子ペーパー、フォントサイズ、高齢化、読みやすさの標準化

## 1 Introduction

Conventional books have a long history of global publication, while the development and spread of e-books has only been present over the two decades. The use of e-book readers has spread remarkably and appears to be expanding for all ages, but questions have arisen regarding their functionality relative to conventional paper, especially, e-paper devices. The readability of liquid crystal displays were fully surveyed till now, but e-paper using e-ink has not been tested on readability.

## 2 Methods

### 2.1 Participants

The participants for this experiment included 99 males and females between the ages of 15 and 76 years (M 46.7, SD 14.7). Participants who usually wore glasses or contact lenses used them for the experiments. We obtained informed consent from all the participants and approval for the study from the Ethical Review Board of the Graduate School of Information Science at Nagoya University.

Table 1 below shows the number of participants divided into four general age groups. The participants were divided into four groups by age: young (29 years and younger), middle-aged (30 to 44 years old), senior middle-aged (45 to 64 years old), and elderly (65 years and older).

Table 1. Participants divided by age groups

| | Young | Middle | Senior Middle | Elderly | Total |
|---|---|---|---|---|---|
| N | 15 | 24 | 44 | 16 | 99 |

We measured the participants' lens cloudiness with an Anterior Eye Segment Analysis System (NIDEK EAS-1000). The typical gradation in the range of cloudiness is from 0 to 250 with the higher number representing greater cloudiness Table 2 below shows the measured values of the readers divided into age groups. The youngest group had sufficient amplitude of lens accommodation. The middle-aged group had sufficient near vision ability, although their accommodation was slightly weaker. The senior middle-aged group had mild presbyopia and problems in near vision work. The elderly group had typical presbyopia and generally had to wear glasses for close vision.

Table 2. Participant lens cloudiness

| Young Mean±S.D | Middle Mean±S.D | Senior Middle Mean±S.D | Elderly Mean±S.D |
|---|---|---|---|
| 46.0±10.0 | 69.1±21.3 | 93.3±21.6 | 148.8±36.6 |

†The range of cloudiness gradation is 0 to 255.

### 2.2 Experimental design

For this experiment, we used three types of reading terminal devices, an e-paper (Amazon Kindle DeluxeTM), a backlit LCD (Apple iPadTM), and as a reference, conventional paper text (with the text printed on PPC paper of 69 % whiteness)[1,2]. Table 3 below shows the shows the general specifications for details each mode of delivery of the text.

Table 3. Device specifications

| | Kindle Deluxe | iPad | Paper |
|---|---|---|---|
| Screen size | 9.7 inch | 9.7 inch | 6 inch |
| Resolution | 150 ppi | 264 ppi | 1200 dpi |

In order to avoid reader bias from brand recognition, the participants read from a bezel covered with white Kent paper with only the screen showing. The test media or test paper was mounted on the center of a board. The text displayed on each medium was set at the same height. The front light level of the KV and the back light level of the iPad were set to maximum levels. The text color was black/dark, and the background color was white/bright.

Figure 1 shows the screen luminance of each device. The screen luminance for the iPad was much higher than for the paper text and Kindle Delux (DX), suggesting that it would perform better outdoors than indoors. We then calculated the contrast ratios from the measured values of the brightness of the background color and text color (Figure 2). The DX showed results on the screen luminance similar to

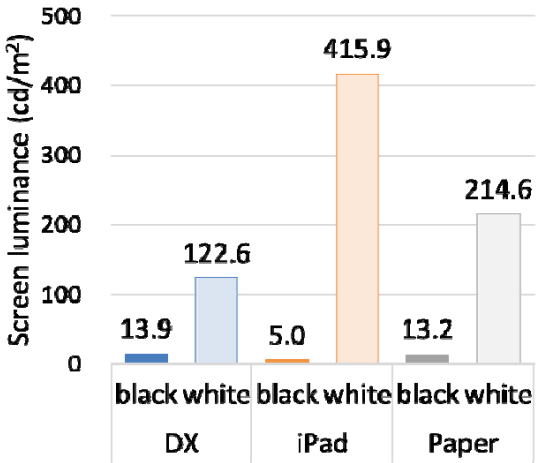


Figure 1. The screen luminance of the black letters on the white background

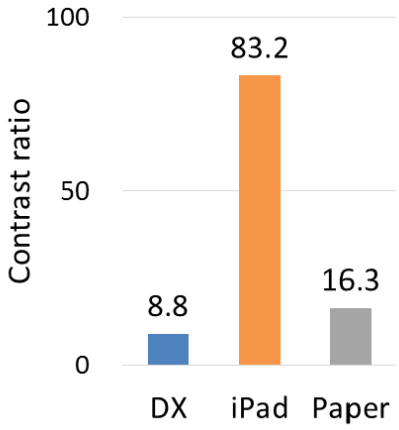


Figure 2. The contrast ratio of each device

conventional paper. The iPad had a higher contrast ratio compared to the other devices because it showed a lower screen luminance of black/dark and a higher screen luminance of white/bright.

We put the reading devices into a compartmental lighting system placed on a desk within a darkened room (Figure 3). We created this lightning system with reference to a previous study[3]. We carried

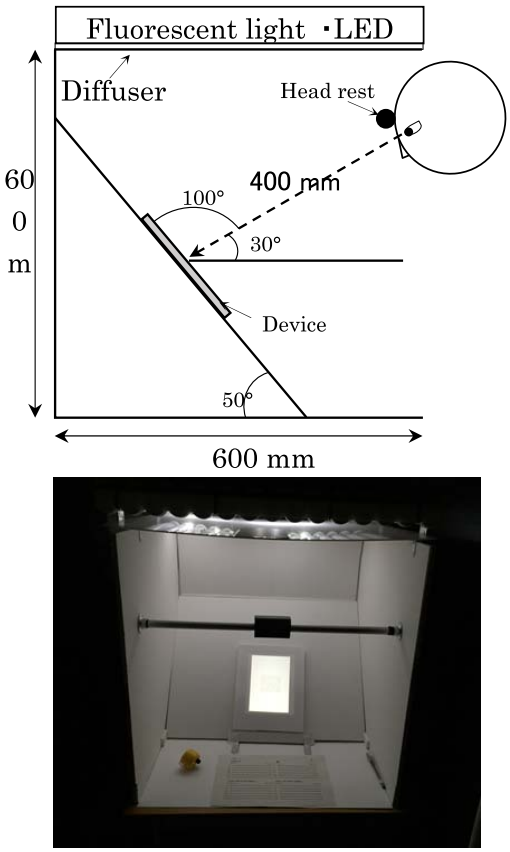


Figure 3. The setup for reading

out the experiment at about the 1,000lx level of illuminance by using a 6500 K LED light source with a fluorescent lamp that had a uniform color temperature. The International Organization of Standards (ISO) recommends an illuminance level of 1,000lx for doing precision work in an office[4]. The headrest for the participant's forehead was kept at a visual distance of 400 mm. The participants looked at the devices at an angle of about 100 degrees to

## 2.2 Task design

The experimental task was a block reading test of a short duration (Figure 4). The display format conformed to that used for evaluation of electronic display devices by the International Organization for Standardization (ISO)[5]. We used random alphanumeric text and unified the size of the characters. The font type was Courier for because it is a uniform monospaced font that has a slight serif ( "bearded" ) which may be of some help to the elderly in reading from e-papers. A monospaced font is one with a fixed width; that is, each letters occupies the same amount of horizontal space[6].

```
LcBs ENckC gplY ELOjPxasTPtAxa
hQ mpU cq VwFdihGamGMztI sj ov
mn KCU QNY hoGqw oS Fa DmsUiLG
BDIqUrGx gLfdaehwd XNj ELdP tL
wYsfmlQa USIg xb BavcTslj QBFv
dK Mb tBH FNcLjY lqZrZN GfrdjC
XSlL NdyTt sa XH wCIIcT rernDq
IEDr vJBpc RvibD LdzY QXb NqEI
kUx zMkm JSxg wmHGF Ltw AOkPjd
BtT xzjT YAI bjYHII AQ vQa Goh
bY nlZ uiVPGzP SZf pH nrzv jDa
tqBBpkRmv QqX AjMGzoJBOrvpKrCZ
Ecwgo zJ ZdyObJn YOz PDBOmzneT
aOe sdNlDL SrysCAcmxbTrMQ sRmA
```

Figure 4. Example of text block for the reading task

As the width of the characters for Courier is uniform, our focus was with the height of the font size. For the first experiment, the font sizes were set at 8 pt (approx. 2.75 mm height), 12 pt (3.25 mm), and 16 pt (5.75 mm). Figure 4 below shows the example of the text blocks for the reading task. The readers performed nine tests (three font sizes per each of the three devices) in which they read the text silently

from the top left and then we measured the total time that it took to complete the procedure.

The participants were asked to count the number of capital letter "*M*" s in the total text with a counter device as they read through the characters. We calculated the correct answer rate from the number "*M*" s that the participants found and the actual number of "*M*" s in the text. The correct answer rate was not entirely exact because the participants sometimes misread other alpha characters as large "*M*" s.

The reading time index and correct answer rate index were divided by the average values of each of the participant's actual values in order to standardize the participant's performance. A high-scoring reading time index meant that it took the participant a long time to read. The high scoring correct answer rate index meant that there was a large number of correct answers.

After each reading, the participants also evaluated the readability of the texts using a Visual Analog Scale (VAS). We converted the values of the VAS into a 100-point scale in the final analysis. Figure 5 provides an example of a subjective evaluation using the VAS method. A high-score on the VAS means a device was very readable. A score of zero (0) as missing values when participants could not read or when the reading time at the answer rate could not be scored.

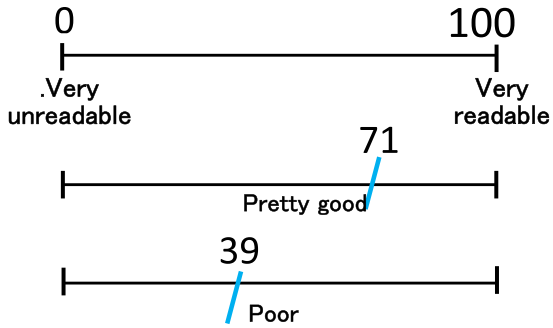


Figure 5. Example of evaluations by VAS

We used a two-way analysis of variance (two-way ANOVA) as a statistical method to evaluate the final results. According to the age grouped analysis, we also used a one-way analysis of variance (ANOVA) and the Scheffé multiple comparison procedure in order to check the differences between the age groups.

## 3 Results

### 3.1 Experiment

Table 4 below shows the results from the two-way ANOVA for participants' evaluation (VAS), reading time index, and correct answer rate index. The VAS showed significant differences among both the types of devices and the font sizes. No two-factor interaction was observed among these fonts. For reading time index, no significant differences were observed among the types of device of the font sizes. The correct answer rate index showed significant differences among the font sizes.

Table 4. Two-way ANOVA for the VAS, reading time index and correct answer rate index

| Source | Df | SS | MS | *F* | |
|---|---|---|---|---|---|
| VAS | | | | | |
| Types of devices(D) | 2 | 7030.47 | 3515.23 | 7.94 | ** |
| Character sizes(C) | 2 | 84091.96 | 42045.98 | 94.96 | ** |
| D×C | 4 | 174.48 | 43.62 | 0.10 | |
| Reading time index | | | | | |
| Types of devices(D) | 2 | 0.004 | 0.002 | 0.226 | |
| Character sizes(C) | 2 | 0.001 | 0.000 | 0.028 | |
| D×C | 4 | 0.045 | 0.011 | 1.163 | |
| Correct answer rate index | | | | | |
| Types of devices(D) | 2 | 0.032 | 0.016 | 0.851 | |
| Character sizes(C) | 2 | 0.330 | 0.165 | 8.732 | ** |
| D×C | 4 | 0.404 | 0.101 | 5.334 | |

† **; p<0.01

Table 5 below shows multiple comparisons between the VAS and correct answer rate index. The VAS scores for 8 pt (app. 2.75 mm) were significantly lower compared to the higher font sizes. The VAS scores for the DX were significantly lower than those of the iPad and Paper. The correct answer rate

Table 5. Multiple comparison for the VAS and correct answer rate index

| | VAS (Mean±S.D.) | | Correct answer rate index (Mean±S.D.) | |
|---|---|---|---|---|
| Character size | | | | |
| 8 pt | 45.9 ± 21.8 | a**, b* | 0.973 ± 0.161 | a**, b** |
| 12 pt | 63.9 ± 19.5 | a** | 1.018 ± 0.122 | a** |
| 16 pt | 68.4 ± 22.0 | b* | 1.009 ± 0.130 | b** |
| Type of device | | | | |
| DX | 55.7 ± 22.9 | a**, b* | | |
| iPad | 62.5 ± 23.3 | a** | | |
| Paper | 60.0 ± 23.1 | b* | | |

† Values in the same column with the same letters are significantly different (**; p<0.01, *; p<0.05)
† For example, 8 pt and 12 pt in the Participants' evaluation column are given the letter 'a'. This means that the subjective evaluation of 12 pt character size is significantly better than that of 8 pt

indexes for the 8 pt size were significantly lower than those of more than 12 pt (3.25 mm) size.

Table 6 below shows a comparison of the VAS by the font sizes for each age group using the one-way ANOVA. The VAS scores for the DX at 12 pt size were significantly lower than for the iPad. Table 7 below shows a comparison of the correct answer rate index by the font sizes for each age group using the one-way ANOVA. The VAS scores for the iPad at 12 pt size were significantly lower than that of the Paper. No significant difference for the reading time index was observed between the font sizes and types of devices. Finally, no significant difference for all indexes were observed among the age groups.

Table 6. Subjects' evaluations using three devices under three character sizes

| | 8 pt (Mean±S.D.) | 12 pt (Mean±S.D.) | 16 pt (Mean±S.D.) |
|---|---|---|---|
| DX | 42.6 ± 22.1 | 60.3 ± 19.0 [a*] | 64.3 ± 21.2 |
| iPad | 49.4 ± 21.0 | 67.1 ± 19.3 [a*] | 71.1 ± 23.5 |
| Pape | 45.8 ± 21.8 | 64.4 ± 19.7 | 69.8 ± 20.5 |

† Values in the same column with the same letters are significantly different (*; p<0.05) using ANOVA & the Scheffé multiple comparison procedure.
† For example, DX and iPad in the 12 pt column are given the letter 'a'. This means that the subjective evaluation of DX device is significantly better than that

Table 7. Multiple comparison for VAS, reading time index and Correct answer rate index

| | 8 pt (Mean±S.D.) | 12 pt (Mean±S.D.) | 16 pt (Mean±S.D.) |
|---|---|---|---|
| DX | 1.014 ± 0.151 | 0.967 ± 0.116 | 1.044 ± 0.101 |
| iPad | 0.936 ± 0.125 [a*] | 1.025 ± 0.071 | 1.058 ± 0.095 [a*] |
| Pape | 0.947 ± 0.173 [a*] | 0.988 ± 0.117 | 1.021 ± 0.100 [a*] |

† Values in the same column with the same letters are significantly different (*; p<0.05) using ANOVA & the Scheffé multiple comparison procedure.

## 4 Discussion

In this study, participants of different ages took a silent reading test to evaluate the readability of three different font sizes used on e-paper device and liquid crystal display with the Courier font type. The aim was to focus on what was the minimum font size that various ages of participants could maintain in their reading performance; that is, testing to see that there was no significant decrease in reading speed and the correct answer rate. In addition, we wished to evaluate to what degree E-paper devices matched the functionality of paper text with respect the Courier font.

In this study, we conducted an experiment under 1,000 lx of illuminance to eliminate the effect of ambient illuminance on the readability of the e-paper and backlit LCD. The ISO also recommends more than a level of 500 lx for indoor reading[4]. In our previous studies, we found no significant differences on the readability of an e-paper, LCD, and paper at the 1,000 lx level[7,8].

We used the Courier font type because it is a monospaced font recommended for such studies by the ISO[6]. In addition, in order to conform to our previous studies, we measured the character heights in millimeters to ensure compatibility. Studies have reported that Courier is a viable font for the elderly suffering with vision problems including macular degeneration[10].

In this experiment, we focused on the e-paper's font features as a display media and how they may effect readability for various ages. We also compared the e-paper and paper with the backlit LCD which are currently available on the market. We referred to previous studies in order to set the experimental text font sizes[10,11]. We set the text font sizes at three levels; 8 pt (character height: 2.75 mm), 12 pt (3.25 mm), 16 pt (5.75 mm).

The reader evaluations showed scores above 45 (on a scale 0 to 100) for each device and font size. In our previous studies, we suggested establishing a baseline score of 45 for VAS to indicate whether participants reached a certain level of proficiency in reading a text[13,14].

The e-paper showed no significant difference compared with other devices for the reading time index and correct answer index. No significant differences for each character size were observed among age groups. All the devices were readable at about the 8 pt size which supports previous studies that looked at visual performance and fatigue[15,16].

## 5 Conclusion

This study investigated the effects of font sizes and aging on the readability of e-papers. We summarize the findings as follows based on this and previous studies and specifically focus on the use of the Courier font size used in each of the two devices.

### 5.1 Font sizes

(1) The screen font size (character height) that participants felt readable was more than 2.75 mm in height (8pt). We recommend a font size of more than 2.75 mm as a default size displayed on the screens of e-papers for all ages.

(2) All groups saw a decrease in reading speed at the 2.75 mm (8 pt) font size so this is why we recommend this size as the minimum limit. The decrease was greater for those over 65 years of age.

### 5.2 Devices

(1) The readability of the two e-books (Amazon Kindle Deluxe and iPad) that we used in this study performed equally to paper text. However, e-books are more advantageous because they allow the reader the opportunity to change the font sizes. This is particularly helpful for the elderly.

(2) The front lighting system (iPad) on e-books works well for improving readability. Furthermore, reading is improved when the screen resolution is greater than 150 ppi as seen with the backlit LCD with the lighting system (iPad).

## References


1) Amazon.com. Accessed November 11. 2016 https://www.amazon.com/Kindle-DX-Wireless-Reader-3G-Global/dp/B002GYWHSQ
2) Apple – iPad, Accessed November 11. 2016 https://support.apple.com/kb/SP647
3) S. Ueki, T. Taguchi, K. Nakamura, Y. Itoh, K. Okamoto, “New Metrics Based on Visual Perception for Evaluating Image Quality,” Proc. IDW ‘06 2006;1: 519-522
4) International Organization for Standardization, “Lighting of indoor work places,” ISO 8995, 2002
5) International Organization for Standardization, “User Performance Test Methods for Electronic Visual displays,” ISO 9241-304, 2008
6) R. Stroud, “Courier Fonts: Everything you wanted to know about courier fonts,” (n.d) (Retrieved October 2, 2016) http://www.rolandstroud.com/downloads/essays/courierfonts.pdf
7) T. Koizuka, S. Sano, T. Kojima, M. Miyao, “Evaluating the Effects of Environmental Illuminance on the Readability of E-books,” SID Symposium Digest of Technical Papers, 2013; 44(1): 571-573
8) S. Matsunami, T. Koizuka, R. P. Lege, T. Kojima, M. Miyao , “The Effects of Ambient Illuminance and Aging on the Evaluation of the Readability of E-paper,” J. Inst. Im. Info. & Tel. Eng., 2015; 69(10): 303-316
10) L. Tarita-Nistor, D. Lam, M. H. Brent., M. J. Steinbach,, E. G. Gonzalez, “Courier: a better font for reading with age-related macular degeneration,” Can. Ophth. Soc., 2013; 48: 56-62
11) D. Lee, Y. Ko, I Shen, C. Chao, “Effect of Light Source, Ambient Illumination, Character Size and Interline Spacing on Visual Performance and Visual Fatigue with Electronic Paper Displays,” Displays, 2011; 32: 1-7
12) D. Lee, K. Shieh, S. Jeng, I. Shen, “Effect of Character Size and Lighting on Legibility of Electronic Papers,” Displays, 2008; 29: 10-17
13) T. Koizuka, Y. Ishii, T. Kojima, N. Ishio, R. P. Lege, M. Miyao, “Proposing a Baseline Setup for Readability using Visual Analog Scale” Proc. IDW ’14, 2014; 21: 1024-1025
14) Y. Ishii, T. Koizuka, R. Cui, T. Kojima, M. Miyao, “Evaluation of Readability for Tablet Devices by the Severity of Cataract Cloudiness,” SID Symposium Digest, 2014; 45(1): 1089–1092
15) S. Sano, T. Kojima and M. Miyao, “The Effect of Illuminance on Visibility during Reading e-books by Age Groups,” Proc. IDW ’12, 2012; 19: 691-694
16) A. Wang,S. Hwang, H. Kuo, S. Jeng, “Effects of ambient illuminance and electronic displays on users’ visual performance for young and elderly users,” J. Soc. Info Display, 2010; 18(9): 629-634
17) Y. Takubo, Y. Hisatake, T Iizuka, T. Kawamura, “Ultra-High Resolution Mobile Displays,” SID

Symposium Digest, 2012; 43(1): 869–872

18) T. Koizuka, Y. Ishii, T. Kojima, R. P. Lege, M. Miyao, “The Contributions of Built-in Light on the Readability in E-paper Devices,” SID Symposium Digest, 2014 45(1): 861-864