THE EFFECT OF FONT SIZE ON THE RETENTION OF TEXT INFORMATION PRESENTED ON COMPUTER DISPLAY AND ON PAPER

Yoshiko MOTOKI

1) Nihon University, Japan

Developing effective computer displays is an important aspect of creating materials for e-learning. This study examined the efficacy of different font sizes used on VDTs (video display terminals) and on print media. It is often thought that people can memorize and remember more when reading from paper than from VDTs. Based on four experiments using different font sizes, the results of the present study showed that different font sizes affect the recall accuracy of content when reading from a VDT. With a 13-point font more content was recalled from a print medium than from a VDT. However, recall differences were not found between the two media for smaller font sizes. This finding suggests that, with font adjustment, the same levels of reading retention can be achieved from VDTs.

Key words: computer display, visual display terminals, font size, reading retention, computer-based learning

Over the past few decades, there has been a steady increase in the use of computers in university settings. Now, in most developed countries, university students are obtaining most of their information and doing a lot of their learning from computers. In Japan, for example, 44.3% of universities are using IT (information technology) and 36.3% of them are introducing e-learning (National Institute of Multimedia Education, 2006). Few research studies, however, have been undertaken on the possible ways in which the actual presentation of information/materials on computer video display terminals (VDTs) affects student learning and retention.

Studies that have looked at the issue of text presentation have tended to be comparative (e.g., comparing visually impaired participants with participants who have normal vision) and have provided little in terms of defining the characteristics of visually presented text information that are most helpful to student learning. For example, Belopolsky and Dubrovsky (1994) pointed out that visually impaired readers are more sensitive to character size and font style than readers with normal vision, and stressed the need to consider font size on computer displays as it affects reading performance at least in some aspects. They therefore developed and evaluated a computer program which

---

I wish to thank Dr. Kazuchika Manabe, Nihon University, under whose supervision this study was conducted. I would like to thank Dr. Emmanuel Manalo, University of Auckland, and Dr. Tadasu Oyama, Nihon University, for their detailed comments and suggestions on the manuscript. I am also grateful to the students of Dokkyo University, who participated in my experiments.

Correspondence concerning this article should be addressed to Yoshiko Motoki, Research Institute of Foreign Language Teaching, Dokkyo University, 1-1, Gakuen-cho, Soka-shi, Saitama 340-0042, Japan (email: motoki@dokkyo.ac.jp).

89
enabled their participants, both with normal vision and with visual impairment, to vary the font size displayed on the screen as they liked. However, Sanocki (1991) had earlier claimed that font size and style changes have more limited effects on reading accuracy. Sanocki conducted three experiments in which font style and size were manipulated. In Experiment 1, the change of font size had less of an effect than the change of style. In Experiments 2 and 3, greater decreases in reading accuracy were observed when only font size was changed than when both font size and style were changed.

More recently, Kiyohara, Nakayama, Kimura, Shimizu, and Shimizu (2003) investigated reading comprehensibility using the respective media of print, CRT (cathode ray tube), and LCD (liquid crystal display). They conducted an experiment in which two task symbols were used, and then compared reading comprehension by using five task texts for each medium. The comprehensibility of the contents of task texts was judged from the results of answers to questions given about the text contents. In this experiment, no significant difference was found between CRT and LCD, but a significant difference was found between printed material and CRT. Kiyohara et al. observed that reading comprehension was greatest with text delivered via print on paper, followed by text delivered via LCD, and least with delivery via CRT display. Their results suggest that comprehension improves when reading from paper. However, no indications were provided about text features, such as font size, that might improve reading from computer screens.

It is important that information displayed on computer VDTs is read easily and retained correctly. Display features are likely to impact on the way in which information is processed, and finding more about this impact can aid developments in the efficacy of computer-based teaching and learning. The present study is therefore geared towards this very purpose. Its aim is to contribute to the development of helpful guidelines for teaching staff who are involved in producing effective and efficient teaching materials.

**EXPERIMENT 1:**

**COMPARISON BETWEEN VDT AND PAPER IN READING THE SAME-SIZED FONT**

In this first experiment, participants’ recall accuracy of content was compared when reading text depicted in the same font size on VDTs and printer paper.

**Method**

*Apparatus:* Four 15-inch liquid crystal displays (COMPAQ TFT5015) were used. The brightness of the stimulus surfaces looked nearly the same, but the luminance measured after the experiment was found to vary among the four experimental seats as follows:

- No. 1: VDT-51.9 cd/m², Paper-43.3 cd/m²
- No. 2: VDT-141 cd/m², Paper-28.8 cd/m²
- No. 3: VDT-101 cd/m², Paper-38.8 cd/m²
- No. 4: VDT-97.2 cd/m², Paper-31.9 cd/m²

The participants were seated in front of the randomly assigned displays in a similar manner. The distances and viewing angles for the sheet of paper and VDT were identical. A sheet of paper was affixed to a liquid crystal display for reading. This was done to control for differences in scrolling and turning pages, as
well as time lags associated with participant experience or inexperience in such tasks. The distance between
the participants and the screen was approximately 45 cm. The arrangement of the sheet and video display is
shown in Fig. 1.

Participants: The participants were 21 liberal arts college students (3 males, 18 females), 18 to 22
years of age (mean age = 20.48 years, SD = 1.29 years), who were using computers on a daily basis to carry
out text-based research and write reports and e-mails. Nineteen of the participants were undergraduates in the
Department of German, while the remaining 2 were from the Department of French. All participants had
normal or corrected visual acuity.

Procedure: Propositional memory was examined. Two kinds of task texts were prepared. Text
selection was based on five criteria: that the text (a) would be of likely interest to students, (b) would have a
similar number of propositions for the measure of recall rate, (c) was originally displayed in Japanese, (d) was
not written in a literary style, and (e) would be too long for participants to memorize all the propositions
during the period allocated for reading. All of the task texts were presented to the students for the first time.
Fig. 2, displays one of the task texts. There were 138 propositions in total in the text given in Task 1 (from
Otake, 2002), and 142 propositions in the text given in Task 2 (from Oishi, 2002).

The texts were prepared in Microsoft Word 2002, using an A4-size (297 mm × 210 mm) horizontal
screen, with a left margin of 38.9 mm, a right margin of 45.1 mm, 13-point Mincho font, 11.6-point kerning,
52 letters per line, and an 18-point line feed interval. The Mincho font used here is the standard font used in
Japan. Before the task text was displayed on the screen, it was printed on A4-size (297 mm × 210 mm) paper.
Comparing the text displayed on the screen and printed on the paper, however, it was found that the size and
distance of letters were physically different (text size on the VDT was 90% of that on printed paper). This
difference may have been caused by some technical differences between the display and the printer.
Therefore, an adjustment was made to render their sizes physically identical. The same adjustment was also
carried out in Experiment 4, where the same font size was again used for both media.

The participants were first divided randomly into two groups, A and B. Group A performed Task 1 on
VDTs and Task 2 on paper. Conversely, Group B performed Task 1 on paper and Task 2 on VDTs. In
addition, each group was subdivided, with sub-groups, A-1 and A-2, completing the two tasks (with the
 corresponding texts) in the opposite order. Group B was divided similarly into sub-groups to control for
possible extraneous task sequencing effects stemming from practice and fatigue.

The experiment took place between 26 November and 17 December 2002. Groups A-1 and A-2 had
six participants each; Groups B-1 and B-2 had four and five participants, respectively. In most cases, the
participants took part in the experiment in groups of four, using the four experimental seats that had been set up.
After reading the content of Task 1 for 2 minutes and 30 seconds, the participants were given 1 minute and 30 seconds to answer Questionnaire 1. After that, they were given 6 minutes to recall and write down the content of as many propositions as possible from the text they had read in Task 1. Next, the task and media were changed. Task 2 was displayed for 2 minutes and 30 seconds, and participants were given 1 minute and 30 seconds to answer Questionnaire 2. They were then allotted 6 minutes for the written recall task. Exact recall of text was not required. However, participants were asked to write down as many significant details of the content as possible.

Questionnaire 1 asked for the participant’s age and evaluated on six-point scales the frequency of computer use for the purposes of playing games and otherwise, as well as the frequency of computer-mobile phone and e-mail use. The purpose was to determine the participants’ familiarity with computers. Questionnaire 2 was designed to evaluate, again on six-point scales, the participant’s history of computer use; amount of time spent reading books, magazines and newspapers; frequency of letter writing; and number of hours spent watching television. The purpose was to determine the volume of printed material that participants were used to reading in a month. Together, these two questionnaires not only assessed the level
of participants’ computer use and the amount of printed material viewed during their daily lives, but they also provided a uniform task to prevent mental text rehearsal during the imposed delay period between text reading and recall.

Instructions concerning cautionary points and time limits were prepared and presented to each group in a uniform fashion. The detailed instructions were: 1) read the sentences for 2 minutes and 30 seconds and memorize the contents (what was written in the text); 2) answer the questions given within 1 minute and 30 seconds; 3) write down the remembered content of the text read within 6 minutes; 4) it is not necessary to reproduce the original text exactly in a word-per-word fashion; 5) simply write down what you can recall of the contents in your own words. These instructions were based on those used by Schmalhofer & Glavanov (1986).

Measures: The accuracy of text recall, expressed as a percentage, was used to measure how much content was retained correctly. First, the propositional content was categorized, based on Anderson’s (1980) propositional network depiction method.

For each of the task texts used in this experiment, the following provisional network was constructed. For example, the opening sentence of Task 1 text was, “I have liked traveling alone, since a long time ago.” In this sentence, the subject, verb and adjective were rendered as separate propositions. Then, ellipses were drawn around each proposition. This example sentence was numbered temporarily as “No. 1”, since it was the first sentence. Then, an arrow (link) labeled “relation” was drawn from each word located in an ellipse, which is called a “node” indicating the concept in memory. Fig. 3 shows the propositional network for this example sentence, with the content within each ellipse as a “proposition”.

Recall accuracy was then determined by calculating the percentage of propositions recalled correctly from the total number possible. Participants were asked to write their responses on paper to eliminate possible variations due to differences in keyboard fluency. There were 138 propositions in total in Task 1 and 143 propositions in Task 2.

The propositions were divided as follows (using the same example sentence for illustration).

Original sentence:
“I have liked traveling alone, since a long time ago.”

Propositions:
<table>
<thead>
<tr>
<th>have liked</th>
<th>alone</th>
<th>traveling</th>
<th>since a long time ago</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Five propositions in total
Recall ratios were calculated as follows:

Response example 1:
Since a long time ago 1 have liked 2 to travel 3
Four out of five recalled in total

Response example 2:
1 like 2 to travel 3
Three out of five recalled in total

Results and Discussion

Fig. 4 shows the participants’ percentages of correct recall of the text propositions according to the display media used. The mean values from the two media were significantly different, $t(20) = 2.5$, $p < .03$. The results indicate that more propositions were recalled from the paper display than from the VDT.

As shown in Fig. 4, the percentage of correct recall of propositions was significantly greater for text read from paper than from a VDT under similar conditions. Thus it may be concluded that recall is better when Japanese text is read from a print medium than from a computer screen at that particular font size.

**Experiment 2: Comparison Between Font Sizes on the VDT**

In Experiment 1, the participants remembered content from hard copy better than from a VDT, when text was displayed under nearly identical conditions (i.e., same font character size and style, viewing distance, view angle, etc.). In Experiment 2, the effect of differences in font size on the ability to recall propositional content was examined in order to determine the most suitable VDT font size for remembering the content of text.
Method

Apparatus: The same display was used as in Experiment 1. However, this time, a 12-point font and a 13.5-point font were used on the VDTs.

Participants: The participants were 11 liberal arts college students (2 males, 9 females), 19 to 26 years old (mean age = 21.55 years, \( SD = 1.97 \) years), none of whom had taken part in Experiment 1. All participants had normal or corrected visual acuity.

Procedure: The texts were the same as those used in Experiment 1. A VDT with a 12-point font was set up as in Experiment 1, using an A4-size (297 mm × 210 mm) monitor, 38.9 mm left margin, 60.0 mm right margin, the standard Mincho font, 9.85-point kerning, 57 characters per line, and an 11.4-point line feed interval. For the 13.5-point font, the monitor displayed a left margin of 38.9 mm and right margin of 28.5 mm, Mincho font, 11.4-point kerning, 57 characters per line, and a 15.55-point line feed interval. As the right margins were different, the text was displayed so as to show the middle of each sentence in the middle of the screen, and the position of line feed, line numbers, and numbers of words on each line were identical.

The participants were again divided into two groups, A and B. Group A completed Task 1 with a 12-point font and Task 2 with a 13.5-font. Conversely, Group B completed Task 1 with a 13.5-point font and Task 2 with a 12-point font. In addition, each group was further divided into sub-groups, such as A-1 and A-2, so that the order in which A-1 and A-2 completed their respective tasks (with the corresponding texts) was reversed. The same protocol was implemented for Group B.

The experiment took place from 12 to 26 July 2003. Groups A-1, A-2, and B-1 consisted of three participants each, while Group B-2 consisted of two participants. Other methodological details were the same as in Experiment 1.

Results and Discussion

Fig. 5 shows the percentage of correct recall of propositions depending on font size. The mean values of the percentage of correct recall in the 12-point font size condition and the 13.5-point font size condition were significantly different, \( t(10) = 2.5, p < .04 \).

As shown in Fig. 5, the percentage of correct recall of propositions was higher with a 12-point than a 13.5-point font. Thus it may be concluded that Japanese sentences displayed on a VDT are recalled more effectively when presented in a 12-point rather than a 13.5-point font size.
EXPERIMENT 3:
COMPARISON BETWEEN VDT AND PAPER IN READING DIFFERENT-SIZED.Fonts

In Experiment 1, when text was displayed on a VDT and on paper under similar conditions, participants were better able to remember the content read from the paper. However, a close examination of the data gathered from both Experiment 1 and Experiment 2 revealed that the percentage of correct recall of proposition when using a 13-point font size on paper (32.4%, from Experiment 1) was identical to that of a 12-point font on a VDT (also 32.4%, from Experiment 2). Therefore, the next experiment was carried out in which the font size was changed for both the VDT and paper displays. Although the texts used were different from those in Experiment 1, the same selection criteria were employed. There were 109 propositions in total in Task 3, and 106 propositions in Task 4.

Method

Apparatus: The same display and methods of preparing the VDTs and paper were used as in Experiment 1, but the letter size used on both VDTs and paper was changed.

Participants: The participants were 16 liberal arts college students (3 males, 13 females), 19 to 24 years old (mean age = 20.88 years, SD = 1.63 years), none of whom had participated in Experiments 1 or 2. All participants had normal or corrected visual acuity.

Procedure: The task texts were displayed on an A4-size (297 mm × 210 mm) horizontal VDT with a 38.9 mm left margin, 45.1 mm right margin, 12-point Mincho font, 10.5-point kerning, 57 characters per line, and an 18-point line feed interval. The same sentences were also prepared on A4-size (297 mm × 210 mm) paper, printed horizontally, with 38.9 mm left margins, 45.1 mm right margins, 13-point Mincho font, 10.5-point kerning, and an 18-point line feed interval.

The participants were divided randomly into Groups A and B. Group A completed Task 3 (with text from Tabinikkishakurabu, 2003) using a VDT with a 12-point font size, and Task 4 (with text from Oishi, 2003) using print on paper with a 13-point font size. Conversely, Group B completed Task 3 using a 13-point font size on paper first, and then Task 4 using a 12-point font size on a VDT. In addition, each group was further subdivided, so that the sequence in which A-1 and A-2 completed the tasks was reversed. The same protocol was implemented for Group B.

The experiment took place between 2 and 10 October 2003. The 16 participants were randomly assigned to the four sub-groups, each of which had four members.

Results and Discussion

Fig. 6 shows the percentage of correct recall of propositions by difference in media. The mean values of the percentage of correct recall of propositions were not significantly different, \( t(15) = 0.17, \text{ ns} \). This finding indicates that with adjustments to the font size used on VDTs and paper, virtually the same percentage of correct recall of propositions for both media can be achieved.

It may be concluded from this finding that when font sizes are altered, the recall of sentences written in Japanese script characters displayed on a VDT can be the same as for sentences printed on paper. The reason for this may be that sentences grouped together in a slightly smaller font are easier to read and more readily segregated from a background screen than those presented in a slightly larger font and filling a larger area of a computer screen.
In Experiment 1, the same 13-point font size was compared for VDTs and print media, while a 12-point font size for VDTs and a 13-point font size for print media were compared in Experiment 3. In Experiment 4, a 10.5-point font size generally used in Japanese text was compared between VDTs and print on paper.

The task texts used in Experiment 4 were the same as those used in Experiment 1.

**Method**

**Apparatus:** The same display as in Experiments 1 through 3 was used and the VDT and paper were prepared as detailed in Experiment 1.

**Participants:** The participants were 12 liberal arts college students (6 males, 6 females), 19 to 27 years old (mean age = 21.58 years, SD = 2.11 years), none of whom had participated in Experiment 1 or 2, although one male and one female had participated in Experiment 3. All participants had normal or corrected visual acuity.

**Procedure:** The task texts were prepared with Microsoft Word 2002 and displayed on an A4 size horizontal monitor with a 38.9 mm left margin, 45.1 mm right margin, 10.5-point Mincho font, 10.5-point kerning, 57 characters per line, and an 18-point line feed interval. The texts were also printed onto paper before being displayed on the screen, and rendered so as to approximate the screen display as closely as possible. The middle of the sentences was aligned with the center of the screen.

Again, the participants were divided into Groups A and B. Group A completed Task 5 using the VDTs and Task 6 using paper. Conversely, Group B completed Task 6 using the VDTs and Task 5 using paper. In addition, each group was further subdivided so that the sequence in which A-1 and A-2 completed the task was reversed. The same protocol was implemented for Group B.

The experiment took place from 4 November to 16 December 2004. Groups A-1, A-2, B-1, and B-2 had three participants each.

**Results and Discussion**

Fig. 7 shows the percentage of correct recall of propositions according to the media used. The mean values of percentage of correct recall of propositions for the two forms of
media used were not significantly different, \( t(12) = 1.23, \) ns.

In Experiment 1, the percentage of correct recall of proposition was higher when a 13-point font size was used on paper. When a smaller 10.5-point font size was used for both paper and VDTs, the significant difference found in Experiment 1 was no longer observed. This finding suggests that with the use of the smaller font size, retention of text information presented on VDTs was improved.

Fig. 8 summarizes the results of Experiments 1, 3, and 4. Results from the VDTs and paper presentations are distinguished by position on the horizontal axis, while the percentages of propositions correctly recalled are shown on the vertical axis.

There was a significant interaction between the media and font-size, \( F(2, 46) = 3.68, p < .05 \). However, the main effect of the media was not significant, \( F(1, 46) = .13, \) ns. This finding suggests that the medium used for presentation on its own is not the critical factor as far as retention of text information is concerned: if font size is appropriately adjusted according to the media used, students can remember text information as effectively irrespective of the media used.

**GENERAL DISCUSSION**

The results of this study indicate that it is easier to remember content read from a VDT when a smaller font size is used for presentation. One possible reason is that, with the smaller font size, the material is more grouped together rather than spread across a wider area of the screen. The grouping together of text may facilitate better cohesion of text components, including the propositions contained in sentences. Unlike the meanings of single words or phrases, which can be grasped at a single glance, the content of longer text — such as sentences and paragraphs — are usually remembered better when they are grouped together in meaningful parts because readers can grasp these parts as chunks (Miller, 1956). The font size that facilitates effective grouping or chunking of text
components may vary according to the media used, which in turn may be influenced by what people are used to. When using a computer, for example, the default font size for Japanese script is 10.5-point. As a consequence, people may have simply become accustomed to reading this size font on VDTs. Further investigations are clearly necessary to determine the actual reasons behind the findings of this study.

Previous research findings suggested better learning outcomes when text is read from a paper print-out than from a computer screen. As noted earlier, for example, Kiyohara et al. (2003) found that reading comprehension of text delivered via print on paper was better than comprehension via VDTs. However, the results of the present experiments show that the same level of recall can be obtained for paper and VDTs when the font size is appropriately adjusted. These results are important to bear in mind when preparing e-learning teaching materials, particularly when choosing font size. They also indicate a need for further research in this area. The intended users of e-learning teaching materials are often largely unspecified people varying in demographic dimensions such as age, gender, occupation, ethnicity, and educational background, as well as the learning surroundings they have to operate in. However, the same learning efficiencies are expected of them given the same information presented in the same way via computer displays. This points to the absolute importance of finding out more about display characteristics that facilitate the most effective learning of information presented via computer screens. The present experiments, for example, did not take into account background or font color, both of which present promising paths for further exploration. As Belopolsky and Dubrovsky (1994) pointed out, individuals with visual impairments are very sensitive to the color, size and style of font as well as the background color: these same elements may have an effect also on readers with normal, unimpaired vision.

The present study contributes to the development of helpful guidelines for preparing
e-learning teaching materials. It highlights the importance of attending to and adjusting font size to facilitate effective retention of text information read from VDTs. However, this study examined only text written in Japanese and, as noted earlier, there is a need for further research in the many different facets of this topic area. It is hoped that this paper will help stimulate such research.

REFERENCES


(Manuscript received November 21, 2005; Revision accepted April 28, 2006)