A Study of Reading with a Handheld Computer

Young Seok Lee

Thesis proposal submitted to the
Faculty of the Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Industrial and Systems Engineering

Advisory Committee

Tonya L. Smith-Jackson, Ph.D., Chair
Robert J. Beaton, Ph.D.
Carmel Vaccare, Ph.D.

August 5, 2003
Blacksburg, Virginia

Keywords: Handheld computers, Readability, Electronic Book,
Reading Study, Personal Digital Assistant (PDA), Mental Workload

A Study of Reading with a Handheld Computer

Young Seok Lee

ABSTRACT

As handheld computers are becoming powerful, portable and wireless, they have recently received considerable attention in education because their multi-functionality and mobility can be used for supporting learning activities in various manner. In spite of the potential in education, handheld computers have several limitations. One of the critical limitations is the small screen size. In an attempt to find factors affecting text readability on the small screen of handheld computers, this study, first, reviewed basic reading process and reading model, previous studies examining potential factors affecting text readability on a small handheld screen.

This study investigated the effect of screen orientation and margin on reading performance (reading speed and comprehension) and subjective impressions (ease of reading, mental workload and satisfaction) with a handheld computer. A 2x2 within-subject factorial design was conducted with 16 participants reading texts from a PDA (Compaq iPAQ 3850).

The results of this study indicated that neither screen orientation nor margin had a significant influence on objective reading performance and subjective impression. It was found that there were no prevailing preference of screen orientation and margin in reading texts from a handheld computer screen. It was also found that participants had a favorable reading experience with the handheld computers. Based on results of experiments and post-experiment questionnaire, a set of recommendations was developed to improve reading from small screens of handheld computers.

DEDICATION

To Jiang Nan

My wife who always loves me and sacrifices herself for me

This achievement and myself are yours forever

I love you

To my family in Korea and China

Your dedication has led me to where I am now

This is the first one of repays that I will make for you from now on

Thanks for your love

ACKNOWLEDGEMENT

Acknowledgements and thanks are due to my advisory committee: Dr. Smith-Jackson, Dr. Beaton, and Dr. Vaccare. Thanks for your insightful and instructive guidance. I would like to express special thanks to Dr. Smith-Jackson for sacrificing her sleeping time to give me helpful guidance. You have shown me what mentorship is.

Many people have been helpful to me in the completion of this thesis. First are all the people who stayed up a night proofreading a series of drafts. Jason Saleem, Chuck Perala, Myrna Callison, and Yogesh Bhatkmande, you are the ones who I should express thanks to for your helpful feedback. I have always appreciated the help my HFES friends have provided. They showed me their friendship with participating in my study, traveling and drinking a toast with me, and taking pictures with me. I also want to pay a special tribute to Sang W. Hong, who has been helpful to me as a friend and a mentor.

Finally, I want to thank Lovedia Cole and Teresa Coalson for helping my document to be processed quickly. Without your help, I would not have made this thesis done on right time.

Thanks to all of you again

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

1.1. Background

Handheld computers have received considerable attention and popularity, with information technology devices becoming powerful, portable and wireless. As handheld computers have become affordable recently, more people carry various personal handheld computers in the form of Personal Digital Assistant (PDA), communicator, personal information organizer, etc. Handheld computers have brought many advantages to our lives. One of the advantages is information access. Unlike a stationary desktop computer, armed with mobility and portability, handheld computers enable people to access to the information they need at any time or place.

Handheld computers have recently received considerable attention in education because their portability, mobility and multi-functionality can be used for supporting learning activities in various manner (Fung, Hennessy, & O'Shea, 1998; Hennessy, 2000; Shepard, 2002). Along with the typical Personal Information Manager (PIM) function such as calendar, contact list, and to-do list, handheld computers enhanced with wireless communication capabilities enable students to access to the Internet for various information ranging from personal e-mail to e-journals of Digital library. Understanding the potential of handheld computers in higher education, a number of colleges and universities are experimenting with ways to use handheld computers in education (Deneen, 2001). Considering the trend of handheld computer use in education or industry, there is no longer any doubt that more people are using handheld computers to access information.

1.2. Problem Statement

In spite of the potential for information access, handheld computers have several limitations. One of the critical limitations is the small screen size. The small screen size limits the space available to display texts, and less information is shown on a screen. Consequently, the user is required to interact with the device more frequently in order to get to desired information. The size of handheld computer screens, however, is unlikely to change very much because their small size is a key factor in their popularity. More users want devices that can be easily carried and held in one hand. The small screen size is a key mobility factor and an adverse factor degrading usefulness. However, the rapid growth in popularity of handheld computers in recent years seems to demonstrate that the effect of screen size is less than the one of increased mobility.

Given that the size of handheld computer screens will likely not change, two questions were raised: “What are the factors affecting text readability on the small screen of handheld computers?” “Can we improve text readability on the small computer screen?”. In an attempt to answer these questions, this thesis reviewed basic reading process and reading model, previous studies examining potential factors affecting text readability on a small handheld screen.

CHAPTER 2. LITERATURE REVIEW

2.1. Reading

2.1.1. Reading Definition and Eye Movements

There have been numerous studies conducted by cognitive psychologists to study the cognitive process that occur during normal silent reading of text. While a number of definitions of reading exist, the operational definition that will be used in this thesis was as follows (Rayner & Pollatsek, 1989):

"Reading is the ability to extract visual information from the page and comprehend the meaning of the text" (p.23).

Although reading is an everyday activity taken for granted by most people, it is a highly complex perceptual and cognitive process that involves a series of fixations and saccadic movement across the page (Rayner & Pollatsek, 1989). During fixation, reading contents are kept in foveal region that is about 2 degree of visual angle. Perception and comprehension of reading content occurs during fixation, taking up to 230 millisecond (ms) for fast readers and 330 ms for average readers (Robeck & Wallace, 1990). Saccadic movement is made to bring the fovea into position where next fixation target is, taking up to 40 ms (Robeck & Wallace, 1990). Backward saccades (right to left) are called regressions and are used to review the text already covered for clarification of incomplete retrieval (Just & Carpenter, 1980).

2.1.2. Reading Model

While there have been various models to describe the reading process, they can be primarily characterized to three models; bottom-up, top-down, and interactive model. The basic idea of bottom-up models is that lower-level units (features and letters) feed into higher-level units (letters and words). These units are transformed through a series of stages with little influence from general world knowledge, contextual information, or higher order processing strategies. With increased reading experience, familiarity of particular words results in automaticity, hence familiar words can be perceived more rapidly (Reynolds, 1984). Therefore, it is called data-driven process.

On the other hand, the top-down model stresses that the general world knowledge and contextual information from the passage being read impacts the reading process. The general world knowledge and contextual information from the text being read are used to form hypotheses concerning what will come next during reading. While reading, people are engaged in a cycle which involves the generation of the hypothesis of what will be read next, confirmation of the hypothesis by minimally sampling the visual information on the page, and then the generation of a new hypothesis for the next (Lindsay & Norman, 1972; Rayner & Pollatsek, 1989).

Interactive models are moderate approaches between the two models. The basic idea of interactive models is that readers draw on both top-down and bottom-up information before an interpretation of the text is executed. The most frequently cited interactive model was proposed by Rumelhart (1977). Although there are criticisms that it is not a comprehensive reading model, this model explains how the reading process is affected by contextual information from the passage and the reader's hypothesis generated from readers' knowledge such as syntactical, semantic, orthographical and lexical knowledge (Rayner & Pollatsek, 1989). In conclusion, while there has been extended controversy of reading models among cognitive psychologists, it is generally accepted that both top-down and bottom-up process influence the reading process together, but which process plays more of a role in reading is still controversial.

2.1.3. Reading Type

In terms of reading speed and processing components, Carver (1990) organized the reading process into five basic reading types. These include scanning, skimming, rauding (reading with comprehension), learning, and memorizing. The five types involve different processing components that may be used when people are reading.

Scanning is searching for a target word by looking at each consecutive word of a passage, hence, lexical access is used to recognize each word. This is the fastest reading executed at 600 Wpm (standard length words per minute using six character spaces as a standard length word) by college students. Skimming is searching a prose passage to find two adjacent words whose order has been reversed or transposed, thus, a semantic encoding component as well as lexical access are used to recognize each word and understand what the word means. This is the second fastest reading performed at 450 Wpm by college students. Rauding is looking at each consecutive word of a passage to comprehend the complete thought of each sentence, thus, sentence integration is added to the lexical accessing and the semantic encoding component in order to understand the meaning of each sentence in relation with other sentences of a passage.

In other words, the rauding is a typical and ordinary reading, and it takes about 300 Wpm for college students. Learning is knowing and remembering the information contained in a passage. It is different from the rauding in that idea remembering is added to comprehension. It takes college students about 200 Wpm. Finally, memorizing is not only remembering the idea, but also rehearsing to recall the words, facts, ideas, or thoughts of a passage on a subsequent occasion. Thus, it is the slowest reading type executed at 138 Wpm by college students.

2.2. Readability and Legibility of Text

Readability is normally focused on continuous, meaningful text and refers to the difficulty or ease with which the meaning of text can be understood. Readability is typically measured by reading rate (De Bruijn, De Mul, & Van Oostendorp, 1992; Duchnicky & Kolers, 1983; Elkerton & Williges, 1984; Hansen & Haas, 1988; Reisel & Shneiderman, 1987), reading comprehension (De Bruijn et al., 1992; Dillon, Richardson, & McKnight, 1990; Reisel & Shneiderman, 1987), identification of misspelled words (Gould, Alfaro, Finn, Haupt, & Minuto, 1987), searching for pre-specified letters (Hansen & Haas, 1988), and reading efficiency defined as 'speed x comprehension' (Muter, 1996).

Legibility generally refers to the degree to which text items can be identified and measured normally by identification of single letters or a small array of letters (Mills & Weldon, 1987). Therefore, it is determined by letter parameters such as contrast, font face, letter size, inter-spacing, etc. It is true to say that readability is affected by some of the letter parameters that affect legibility, including luminance, contrast, letter size, and target spacing. However, the performance in legibility does not necessarily reflect performance in readability. For example, although legibility improves as letter size increases, when letters become too large, readability can be adversely affected (Legge, G., Rubin, & Schleske, 1985).

2.3. Factors Affecting Text Readability on Computer Screen

A number of studies have been conducted to investigate readability text from a computer screen. A comprehensive review of empirical literature was given by Dillon (1992) and Mills and Weldon (1987). The results of these studies reported that a number of factors could affect the readability of computer screen.

2.3.1. The Effect of Screen Size

Numerous studies have been conducted to investigate the effect of screen size on reading speed and comprehension, but results of the studies are contradictory. Some studies found a significant effect of screen size on reading performance, but others did not. Table 1 is a summary of those studies. Although the effect of screen size can be broken down into the effect of window height and line length, most studies concentrated on the effect of window height.

Table 1. Summary of Previous Studies on Effect of Screen Size

Experiment	Task	Reading material	Screen size	Measure (* Significance)
Duchnicky and Kolers (1983)	Reading for comprehension	30 noncontiguous passages of about 300 words	Window Height: 1,2,3,4,20 lines Line length: 6.2, 12.5, 18.7cm	Reading rate *, Comprehension
Elkerton and Williges (1984)	search task	45 or 200 line data or text file	Window Height: 1, 7, 13, 19 lines	completion time*
Hansen and Hass (1988)	Reordering a scrambled text	1200 word text	50lines * 90characters & 24lines * 80 characters	Completion time*
Dillon et al. (1990)	Reading a lengthy text for comprehension	a 3500-word lengthy text	Window Height: 20, 60 line	Reading rate, Comprehension, Reading rate x comprehension, subjective impression
De Bruijn and De Mul & Van Oostendorp (1992)	Reading for comprehension with secondary task	Davis Reading Test (5 short text and a total of 23 multiple choice questions)	Window Height: 12 inch (23 line), 15 inch (60 line)	Reading time*, Comprehension, Cognitive effort
Reisel and Shneiderman (1987)	Program reading	Program source (e.g. Dijkstra's Algorithm)	Window Height: 10, 22, 60 and 120 lines	Reading time*, Comprehension*, Preference*
Dyson and Kipping (1998)	Reading for comprehension	800 words text from Microsoft Network	Line Length: 25,40,55,70,85,100 cpl	Reading time*, Comprehension, Preference*

* Significant effect found (p<0.05)

In the literature review, the study by Hansen & Hass (1988) changed the two elements (window height and line length) while using the two different screen sizes (24 lines by 80 characters and 50 lines by 90 characters). A statistically significant difference (p<0.05) of completion time to reordering the scrambled text was found between the large and small windows. The results showed that the large screen allowed a better overview and could have a beneficial influence on the cognitive processing of visually presented text by displaying more information at a time. This idea has been supported by many studies (Buchanan, Farrant, Jones, Marsden, Pazzani, & Thimbleby, 2001; De Bruijn et al., 1992; Hansen & Haas, 1988; Jones, Marsden, Mohd-Nasir, & Boone, 1999).

Window Height

Although a number of studies have investigated the effect of window height, the results of the studies are contradictory. Duchnicky and Kolers (1983) considered the effect of window height on reading constantly scrolling text. Their study reported that text presented in the 4-line window was read as efficiently as text in the 20-line windows, and text in 1- or 2-line windows was read only 9% more slowly than text in 20-line windows.

Elkerton and Williges (1984) reported in their study with 1-, 7-, 13-, and 19-line displays that there were no significant difference in search time between 7-, 13-, and 19-line windows. Dillon et al. (1990) also investigated the effect of window height using 20- and 60-line window and reported that display size had no significant effect on comprehension or reading times.

Contrary to these results, several researchers reported the significant effect of screen size on reading performance. Kruk and Muter (1984) indicated that the limited amount of information on a screen 'page' is a critical factor in reading performance. In the study conducted by Reisel and Shneiderman (1987), four window sizes (10, 22, 60 and 120 lines) were compared for program reading, and a significant effect of window height was found on both reading time and comprehension ($p<0.05$). De Bruijn et al. (1992) investigated the effects of window height and text layout and reported that a significant effect of window height on reading time was found between a 23- line window and a 60-line window. However, neither window height nor text layout has a significant influence on the cognitive effort and comprehension.

In spite of the contradictory results of the previous studies, most of the studies suggested that a large window height produces faster reading speed, but the impact of varying the window height is not dramatic as long as window height is over four lines.

Line Length

The effect of line length on reading was thoroughly investigated by Tinker (1963) using printed text on paper. The result of his study reported that although the optimal line length for text on the printed page varies as a function of both type size and interline spacing, the line length from 8 cm to 13 cm (approximately 52 characters) resulted in generally best reading performance. Results also showed that reading performance was impaired from either shorter or longer lines, since very short lines elicit more and longer fixation, whereas very long lines make it difficult for the eyes to make an accurate return sweep from the end of one line to the beginning of the next.

Duchnicky and Kolers (1983) conducted a study with a computer screen. The study used three different line lengths: full screen width (18.7 cm; 78 "M" letters in upper case), two-thirds width (12.5 cm; 52 letters), and one-third-screen width (6.2 cm; 26 letters). The result of the study found that reading speed increased 28% from the shortest line length to the longest line length. However, no significant difference was found in comprehension since subjects maintained a constant level of comprehension by varying their reading speed. The result of this study was somewhat different from Tinker's study in that the text presented on the full screen (18.7 cm, 78 characters) was read as efficiently as text in the two-third-screen width (12.5 cm, 52 characters). For the discrepancy, Duchnicky and Kolers (1983) reasoned that the upward movement of the scrolled text reduced the difficulties associated with longer lines of text.

Dyson and Kipping (1998) also reported that reading speed increased as characters per line increased. In their study using 25, 40, 55, 70, 85, and 100 characters per line (cpl), the slowest reading speed was found at 25 cpl and the fastest at 100 cpl, but users preferred 55 cpl. Based on the literature review, line length has an effect on reading speed, and its optimal length is approximately 55 cpl for both computer screen and paper. However, it is argued that the physical characteristics of the display, such as abnormal screen configuration with too much blank area, would affect reading performance since their second experiment with modification failed to prove significant effect of line length on reading performance.

According to results of the previous studies, it appears that the optimum line length is approximately 55 cpl for both computer screen and paper. The impact of varying the line length is much more dramatic than window height, and 25 cpl is the critical point to significantly affect reading performance. Given that previous studies reported contradictory results, this issue needs to be revisited.

Screen Orientation

Most computer screens can be adjusted in two ways: portrait and landscape orientation. Portrait orientation has a greater height than a width (e.g. 6cm wide by 8cm high), and landscape has a greater width than a height (e.g. 8cm wide by 6cm high). Most computer screens, including television screens, are landscape-oriented. Few studies have been conducted for the effect of screen orientation. Gould et al. (1987) conducted a study using CRT displays and reported that screen orientation might contribute to slower reading for comprehension reading not for proofreading.

However, the clear preference for portrait orientation was found by Weardon (1998). In his study with about 200 students from various majors, portrait was strongly preferred when they read documents regardless of medium type (i.e. paper or computer).

2.3.2. Margin Width

Although comments on the usefulness of margins have been frequent, only a few empirical studies have been conducted to determine the effect of varying margin widths on readability. Paterson and Tinker (1940) stated the margin must confine the attention repeatedly as a person reads one line after another. Recently, Morris and Hinrich (1996) implied that wide margins make reading easier by adding the visual contrast.

Empirical studies, however, have shown a different result from what most researchers believe. The study conducted by Paterson and Tinker (1940) examined the effect of margins using a reading material with 7/8-inch margins. The speed of reading the material was compared with the speed of reading material without margins. The results for 190 readers showed no significant speed difference in reading with or without margins, and they concluded margins do not promote greater legibility.

Recently, Youngman and Scharff (1998) investigated the influence of text width and margin width on reading (scanning) from a computer screen. Although no significant effect of margin width was found, the larger margin widths produced the fastest reading on the shortest line length (4 inch). They claimed that larger margins with short line length made the text stand out more. While none of the empirical data showed a significant effect of margin width on reading performance, it is generally accepted that margin space is an important factor for readability. Text without margin makes it difficult for the eyes to make an accurate return sweep. Consequently, returning to the current line or jumping ahead to two lines can occur (Siegel, 1997).

Although lack of a margin will not cause a significant time difference in returning to the right line, it requires increased cognitive effort. Thus, text without margin will be less preferred. However, a margin can be a very sensitive issue in the small screen size of handheld computers in terms of information amount shown per page because having a margin width will reduce the line length of the small screen considerably. Therefore it is possible that people will not prefer having margin in a small screen. Given that few studies were conducted to investigate the margin effect on a small screen, an empirical study needs to be done.

2.3.3. Text Presentation

The small screen size on handheld computers limits information amount per page. Most documents will not likely be presented on a single screen. Therefore, handheld computer users must manipulate the document to reach the passages they wish to read. The ways in which text is presented on a computer screen are referred to as text presentation methods (Mills & Weldon, 1987). Text presentation methods can be divided into static and dynamic presentation. The static presentation method presents texts in a stationary mode with a spatial layout in the same manner as a printed page of the paper. The text is read by scanning with the eyes. In the dynamic presentation, which is possible only with computers, text is scanned across the screen in various ways, and eyes are usually fixated.

Static Presentation Methods

The most commonly used types of static presentation are scrolling and paging. Several studies have examined the effects of the two types of text presentation on reading speed and comprehension (Dillon et al., 1990; Dyson & Kipping, 1998; Piolat, Roussey, & Thunin, 1997; Schwarz, Beldie, & Pastoor, 1983). The results of these studies showed that text presentation does not affect reading speed and comprehension, but the preference of paging was found.

Dillon et al. (1990) conducted a study to investigate the effect of split texts. He found readers in the split text condition returned to the previous page twice more often than readers in the non-split text condition. In the study, he claimed that “The basic elements of comprehension in reading are propositions, sentences and paragraphs. Thus, splitting paragraph and sentences across screens is likely to disrupt the process of comprehension by placing an extra burden on the limited capacity of working memory to hold the sense of the current conceptual unit. Therefore, scrolling will be likely to help readers comprehend the text content better than paging, since there is virtually no split across pages with scrolling” (p.216).

Given that people do not like scrolling and the small screen on handheld computers requires more scrolling, it is obvious that paging is preferred on the small screen. However, much more sentences are also split in the paging on the small screen. Therefore, how split texts on a small screen affect reading performance and preference needs to be studied.

Dynamic Presentation Methods

As computers made it possible to move texts on the screen, several attempts have been made to optimize reading from a small screen by having the text move in various ways. The advantage of dynamic presentation is that little interaction is required from the reader. Mills and Weldon (1987) explained the five dynamic presentation methods: vertical scrolling, Times Square scrolling, rapid serial visual presentation (RSVP), segmented text, and controlled-rate text. Many studies (Kang & Muter, 1989; Muter, 1996; Rahman & Muter, 1999) have tested two main methods of dynamic text presentation; RSVP and the Times Square format. Their studies found people with the dynamic presentations read faster than the static presentation at the same comprehension level.

However, the major drawback of dynamic presentations is that they are still unfamiliar reading methods to most people, and they require a high mental workload unless some training is given. This drawback has not been overcome despite several attempts. One of the attempt was an adaptive algorithm in which each text chunk exposure time is adjusted with respect to the text appearing in the RSVP text presentation window to decrease the task load (Öquist, 2001). However, this study also failed to decrease a cognitive workload. The most important point is that while it is obvious the dynamic presentation shows a great advantage on a very small screen (e.g., a cellular telephone), there is no real reason to use it on a handheld computers screen. It is further argued that everyday reading is habituated through a long history, so many people are reluctant to adopt new reading skills.

2.3.4. Typographical Issues

Font Type

The most common font types on a computer screen are Times New Roman and Georgia (serif) and Arial and Verdana (sans serif). Times New Roman serif fonts were designed for both legibility and economy of print space and became a popular font for print and computer screen. Georgia, Arial, and Verdana were designed specifically for computer screen to increase legibility, having a relatively large x-height compared to Times New Roman.

The effect of font type on readability from a computer screen was thoroughly investigated by (Boyarski, Neuwirth, Forlizzi, & Regli, 1998). They examined the readability and subjective preferences of a set of fonts (i.e. Times New Roman, Georgia, and Verdana) for screen display. The result of the study reported that no differences were found in the reading speed and comprehension among the three fonts. However, preference for Georgia over Times New Roman was found. The result of the study conducted by Bernard and Mills (2002) also reported that Arial font was the most preferred and it had a slight advantage over the Times New Roman font for the best font choice. In another study by Bernard et al. (2002) reported that older people greatly preferred sans serif fonts (i.e. Verdana and Arial) to serif font (i.e. Georgia and Times).

The results of the previous studies using large computer screens imply that font type does not significantly affect reading performance in general, but sans serif fonts are preferred to serif fonts. However, sans serif fonts occupy more physical area due to the extra space between characters, so that less number of characters is presented per line, resulting in less number of characters at a page. Considering the small screen size on handheld computers, it may be important to maximize the amount of information on a screen, therefore, it is possible that users prefer serif fonts that result in more density of character per line. However, few studies have investigated the effect of font types on a small screen.

Font Size

In general, the legibility of individual characters increases as character size increases. However, previous studies reported that smaller size characters produced faster reading that larger characters (Mills & Weldon, 1987). This is because larger size characters may require scanning larger areas in order to obtain the same amount of information.

Again, considering the small screen size of handheld computers, font size might be a sensitive factor affecting readability of text on a small screen. Preference of font size on a small screen might be different from a desktop computer. Nonetheless, few studies have been conducted.

Anti-aliasing

Most computer displays are raster displays, typically containing dot-matrix characters and lines that give the appearance of 'staircasing'. By smoothing out the jagged edges of characters, anti-aliasing makes the text more legible. There are two types of anti-aliasing: grayscale anti-aliasing and color anti-aliasing. Grayscale anti-aliasing attempts to give the impression of smoothness to the shapes of bitmapped characters. This is done by using a range of gray pixels to fill in the jagged gaps around the edges of letterforms. Although it makes text easy to read, it has a critical drawback in that it makes text look fuzzy and blurred since it can only 'gray out' whole pixels. On the other hand, color anti-aliasing uses the red, green, and blue components of each pixel. By controlling the intensity of these sub-pixels, color anti-aliasing can increase a display's horizontal resolution three times, which results in less blurred text (Felici, 2000).

Nevertheless, previous studies have reported that anti-aliased text does not affect reading performance. According to the experiment with anti-aliased text by Gould et al (1987), reading anti-aliased characters was not significantly different from reading dot-matrix characters in terms of reading speed. However, many participants preferred the anti-aliased characters. Supporting these findings, the study done by Boyarski (1998) reported participants rated anti-aliased fonts as more legible than dot-matrix fonts, though no significant effect was found.

However, the study by Bernard (2002) in which font type, size, and anti-aliasing were considered together, reported that while anti-aliasing didn't affect reading performance, regardless of font type, dot-matrix fonts were preferred over the anti-aliased fonts in terms of reading experience and preference. Bernard claimed that anti-aliasing tended to blur the text to some extent, thus possibly making it less legible for specific types and sizes of fonts than dot matrix fonts. Warren (2002) also mentioned that gray anti-aliasing tends to blur the smaller font size because of pixels used for anti-aliasing outnumbers those for original text. Therefore, it is expected that gray anti-aliasing will make texts on a handheld computer screen less legible than dot-matrix fonts. Consequently, dot-matrix fonts will be preferred over the gray anti-aliasing. While previous studies indicated that anti-aliasing does not directly affect reading performance, given that preference influences long-term use, a study needs to be done with handheld computers.

2.4. Measures of Reading

To analyze reading performance, various measures were used in the previous studies. A comprehensive review of empirical literature was given by Dillon (1992). Dillon classified the assessment of reading activity into two types of measure: outcome measures and process measures. This thesis followed his classification. Outcome measures of reading activity focus on speed, accuracy, fatigue, comprehension, and preference. Process measures concentrate on how people read texts.

2.4.1. Outcome Measures

Reading speed is the most common measure in reading study. It is measured using stopwatch or computer itself when tasks are done with computers (Gould et al., 1987; Kak, 1981; Muter, Latremouille, Treurniet, & Beam, 1982). Reading comprehension is another common measure, though how to measure comprehension is still debatable.

One way to measure comprehension is done by asking readers a number of questions about the text they read, but it is not clear that whether it is measuring memory or comprehension (Dillon, 1992). However, as Rahman and Muter (1999) did in their study, using passages and questions from the comprehension portion of the Graduate Record Examination study guide seems to be reasonable since it is a validated test by a number of students and its questions inquire more about implication of the context rather than certain facts in the passage.

Another way was done by Van Dijk and Kinstsch (1983) , Egan et al. (1989), and McKnight et al. (1992). They assessed comprehension by analyzing reader's written essays. It seems valid in that it conforms to the type of assessment of usually employed in colleges, but it has a drawback of insensitivity that can not reflect subtle differences (Dillon, 1992).

Effective Reading speed is another measure. It is calculated by reading speed times percent correct on a comprehension test to investigate a relationship between reading speed and comprehension (Muter & Maurutto, 1991). Subjective mental workload was another measure used in reading studies (Muter, 1996; Oquist, 2001). Many studies examining dynamic presentation methods used this measure to assess mental workload during reading tasks. The workload was defined by Hart and Staveland (1988) as follows: "a hypothetical construct that represents the cost incurred by a human operator to achieve a particular level of performance"(p.140). Although there are many methods to assess subjective mental workload such as the Cooper-Harper and Subjective Workload Assessment Technique (SWAT), National Aeronautical Space Administration Task Load Index (NASA-TLX) has been commonly used in reading studies. Please refer to Hart & Staveland (1988) for NASA-TLX.

2.4.2. Process Measures

While techniques like measuring eye movement have been devised, it is still unclear how to obtain accurate process data (Dillon, 1992). Although eye movement record has significantly contributed to reading theory development, it has not been supported as a suitable method in that such equipment is intrusive, often requiring the reader to remain immobile through the reading. Also, it is not clear whether eye movement reflect what the reader think or try to do. Verbal protocols asking the reader to "think aloud" have been effectively used by researchers to gain information on reading strategies in that they are cheap, relatively naturalistic and physically non-intrusive (Olshavsky, 1977). However, this technique has been criticized for interfering with the normal cognitive processing (Ericsson & Simon, 1993; Nisbett & Wilson, 1977). Meanwhile, diary study, structured interview, and observations have been used in recent publications (O'Hara & Sellen, 1997). They appear to be an effective method as long as the interview is organized well enough to elicit information from the subjects.

CHAPTER 3. RESEARCH OBJECTIVES

The literature review identified a number of potential factors affecting text readability on a handheld computer screen. They included line length, window height, screen orientation, margin, text presentation and typographical issues such as font type, font size and anti-aliasing. Out of many factors, this study investigated the effect of two factors that were hypothesized to be sensitive on text readability on a handheld computer screen, when all other factors are held constant. These two factors are screen orientation and margin and were selected for the following reasons.

Firstly, changing screen orientation can be considered as a change on two parameters of screen size: line length and window height. The landscape orientation has increased line length and decreased window height, compared to the portrait orientation. Given that line length is a critical factor for text readability, it was hypothesized that landscape orientation improves reading performance compared to the portrait. Secondly, screen orientation is a determining form factor in hardware design of handheld computers. The small size of handheld computers restricts their screen size to that which can be held in a hand. The limit to which line length can be extended in a handheld computer is about the line length of landscape mode. Thirdly, from the literature review, the effect of margin was not significant in the case of a desktop computer . Having a margin on a small screen reduces a higher portion of line length as compared to a desktop computer. Consequently, the line length is reduced again. Given that line length is a critical factor for text readability, it was hypothesized that no margin would produce better reading performance in a handheld computer.

In summary, the two objectives of this study are:

1) To investigate the effect of screen orientation and margin on reading performance and subjective impressions with handheld computers
2) To investigate preference of these two factors in a handheld computer.

CHAPTER 4. METHOD

4.1. Experimental Design

A full factorial, 2x2 within-subject design was used in this experiment. The two within-subject factors include the screen orientation (O) and margin (M). Each factor had two levels. The levels for orientation included portrait (Horizontal length 6.13 cm x Height 8.17cm) and landscape (reversed horizontal length and height). The levels for margin included 'with margin' and 'without margin'. A 5mm margin was provided horizontally on both side for 'with margin' level. Two levels of factors created four treatment conditions as shown in Table 2.

Table 2. Overview of Experimental Condition

		Margin (M)	
		With Margin	Without Margin
Orientation (O)	Portrait	Condition A ($S_1 - S_{16}$)	Condition B ($S_1 - S_{16}$)
	Landscape	Condition C ($S_1 - S_{16}$)	Condition D ($S_1 - S_{16}$)

Each treatment condition had different line length and number of lines. Without margin, the portrait mode presented an average of 42 characters across the 16 lines. In the landscape mode, an average of 55 characters were presented across the 10 lines. With 5 mm margin taking space for approximately 6 characters per line, 36 characters were displayed across 16 lines in the portrait mode, and 49 characters across 10 lines for the landscape mode. Table 3 shows the line length and window height for each treatment condition.

Because of the nature of the reading task, it was essential to employ a within-subject design in which effects of the experimental factors were assessed by participants performing the same task in all experimental conditions. Since it was impossible to have participants read the same passage on all conditions, four different passages with similar difficulty level were used in this experiment (Refer to reading material section). In order to remove the carryover effect that can cause biased experimental result in the within-subject design, the order of treatment conditions and reading materials were counterbalanced across participants using a balanced Latin square. This ensured that each passage was applied to all treatment conditions in different orders. The pairing of reading passages and experimental conditions was shown in Table 4 along with its order. Please refer to the reading material section for the detailed information of reading passages.

Table 3. Line Length and Window Height for Each Treatment Condition

	Experimental Treatment Condition			
Treatment	A	B	C	D
Line Length	36 cpl	42 cpl	49 cpl	55 cpl
Window Height	16 lines	10 lines	16 lines	10 lines

*cpl (characters per line): the number of "x" characters in lower case in a size of 10 point of Tahoma font.

Table 4. Reading Passages and Experimental Conditions Counterbalanced

	Treatment Order Ex) 1 A : Passage 1 in the condition A							
Subject	1st		2nd		3rd		4th	
S1	1	A	2	B	4	D	3	C
S2	2	B	3	C	1	A	4	D
S3	3	C	4	D	2	B	1	A
S4	4	D	1	A	3	C	2	B
S5	2	A	3	B	1	D	4	C
S6	3	B	4	C	2	A	1	D
S7	4	C	1	D	3	B	2	A
S8	1	D	2	A	4	C	3	B
S9	3	A	4	B	2	D	1	C
S10	4	B	1	C	3	A	2	D
S11	1	C	2	D	4	B	3	A
S12	2	D	3	A	1	C	4	B
S13	4	A	1	B	3	D	2	C
S14	1	B	2	C	4	A	3	D
S15	2	C	3	D	1	B	4	A
S16	3	D	4	A	2	C	1	B

4.2. Dependent Measures

The dependent measures collected during this experiment were classified as objective and subjective measures. The objective measures included reading speed, reading comprehension, and reading proficiency. The subjective measures included perception of ease of reading, subjective mental workload, and overall satisfaction. Preference of orientation, margin, and combination of the two factors was collected in the post-experiment interview.

4.2.1. Objective Measures

1) **Reading speed:** The time taken to read each passage was recorded in words per minute (wpm).
2) **Reading comprehension:** On completion of the reading task, participants were given five multiple choice type comprehension questions. The percentage of correct answers out of five was recorded.
3) **Reading proficiency (Reading speed x Reading comprehension)** was calculated as an overall reading performance index. This was used to investigate a possible relationship between reading speed and comprehension.

4.2.2. Subjective Measures

1) **Perception of ease of reading:** This was to measure subjective opinion on ease of reading and was rated using the scale shown in Figure 1. Participants were asked to mark anywhere with numerical value.

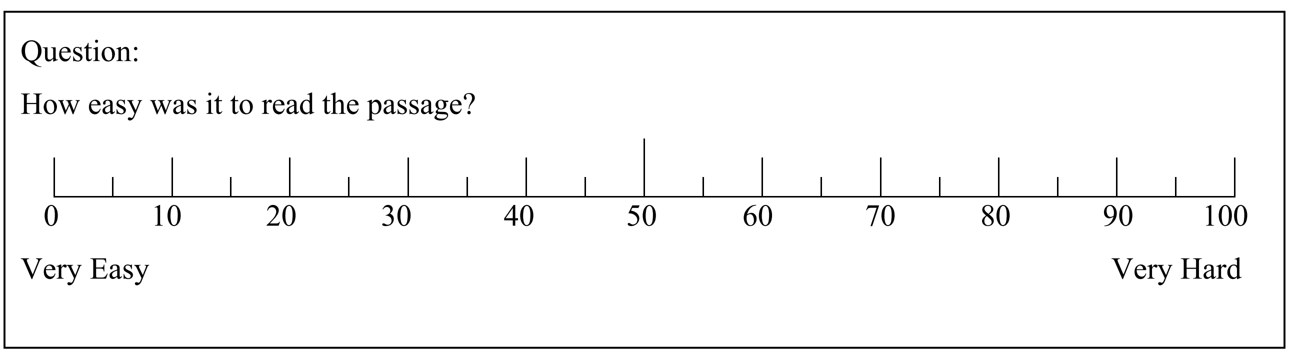


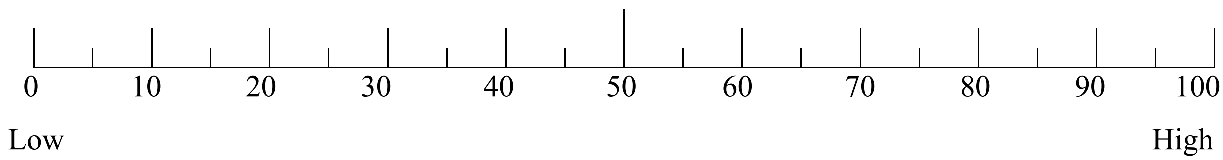
Figure 1. Rating Scale for Subjective Perception of Ease of Reading

2) **Subjective Mental Workload:** Subjective mental workload was measured because this is an important factor in reading from a small screen. Texts on a small screen should be presented to minimize subjective mental workload for efficient reading. The NASA-TLX (Hart & Staveland, 1988) was used to measure the subjective mental workload. Although this method provides six

dimensions, this experiment used only four dimensions (i.e. mental demand, performance, effort, and frustration) since the two dimensions (i.e. physical demand and temporal demand) are not related to the reading task. Figure 2 shows the rating scale of the NASA-TLX.

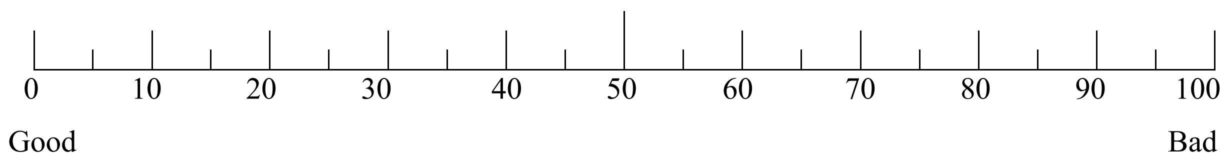
1) **Mental Demand**

How much mental and perceptual activity was required?



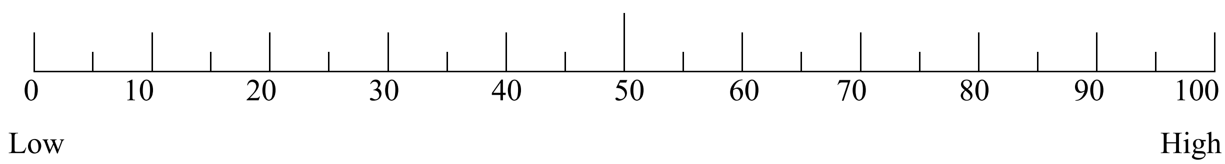
2) **Performance**

How satisfied were you with your reading performance?



3) **Effort**

How hard did you have to work to accomplish your level of reading performance?



4) **Frustration**

How insecure, discouraged, irritated, and stressed versus secure, gratified, content, and relaxed did you feel during the reading task?

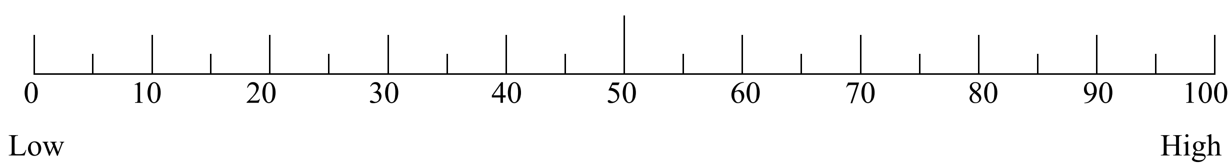


Figure 2. Four Dimensions of The NASA-TLX Used in The Experiment

3) **Overall Satisfaction:** The overall satisfaction for each experimental condition was measured using the rating scale shown in Figure 3.

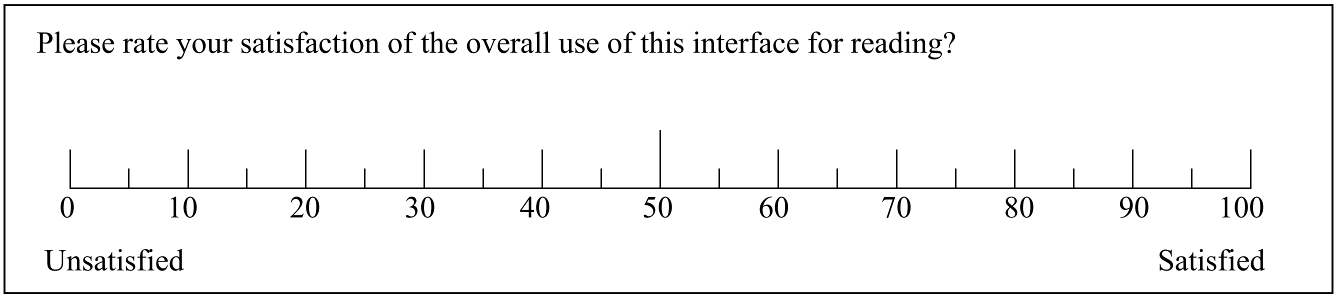


Figure 3. Rating Scale for The Overall Satisfaction

4) **Preference:** The preference of orientation, margin, and combination of the two factors were measured from the post-experiment questionnaire shown in Appendix D. This preference measurement was made to check the consistency of participants' rating by comparing the overall satisfaction score.

4.3. Experimental Hypothesis

The experimental hypothesis for this experiment is listed below:

There will be a main effect of screen orientation, margin, and interaction between the two factors on reading speed, reading proficiency, perception of ease of reading, mental workload, satisfaction, and preference.

4.4. Participants

Sixteen university students participated in the experiment. They were recruited through advertisements within Virginia Tech. Participants were screened according to the following criteria:

(1) Native English speakers whose official language is English (e.g. American, Indian, Canadian, etc.)
(2) Appropriate reading skills to pass the screening test. (See procedure)
(3) Were at least 18 years of age

It was essential to recruit native English speakers to avoid bias caused by vocabulary limitation. The initial screening was conducted via information in the advertisement, and participants were further screened using a sample reading test (See procedure for the sample reading test). There were nine female (56%) and seven male (44%) students. Three of 16 participants were undergraduate students, and the rest were graduate students. Participants ranged in age from 22 to 46 (*M*=26.56, *SD*=5.45). Eight participants were Personal Digital Assistant (PDA) owners. The complete participants' demographic data can be found in Appendix E. Each participant was presented with experimental conditions and passages in different orders, following the counterbalanced order shown in Table 5 of the experimental condition section.

4.5. Equipment

The experiment was conducted using a Compaq iPAQ 3850 PDA (Appendix B). The Compaq iPAQ 3850 has a 3.8-inch screen (8.17cm x 6.13cm) with a resolution of 320 x 240 pixels. The PDA was connected to the Toshiba Satellite 2800 Laptop computer via Universal Serial Bus (USB) cable. This was done so the experimenter could observe all actions made by participants while reading passages. The Remote Display Control application was used to display the actions on the PDA to the laptop computer.

4.6. Reading Material

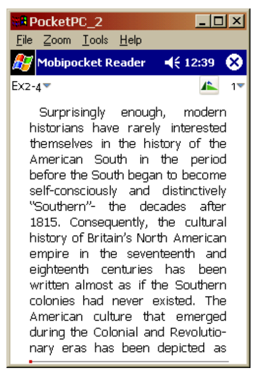
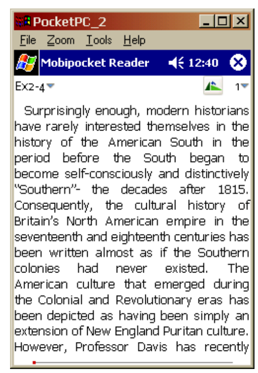
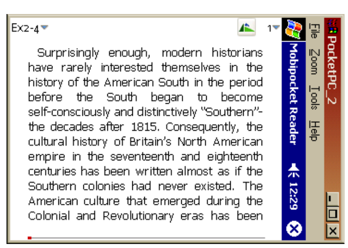
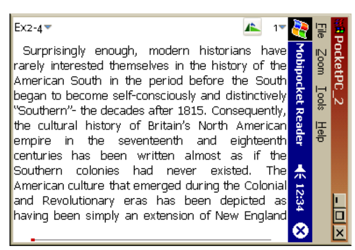
A total of four passages were selected from the comprehension part of the Graduate Record Examination study guide for the General Test (ETS, 2002). In an attempt to minimize the difference of difficulty level, the reading passages were excerpted from the publicly validated test. The passages had a mean of 452 words with a standard deviation of 19 words, and the words ranged from 433 to 484.

The passages were displayed in the Mobipocket reader for Pocket PC™ in different modes of the experimental factors (orientation and margin) as explained in the experimental design section. The passages were displayed in left justification using 10 point Tahoma font with 12 point interlinear spacing. The text was displayed in black color and ClearType™ color anti-aliasing font technique. Table 5 shows the number of words of the passages, and Table 6 illustrates an example of the passage display in each experimental treatment condition.

Table 5. An Overview of Reading Materials

Index	Words
P1	484
P2	449
P3	433
P4	443

Table 6. Text Display Used in Each Experimental Conditions

<table>
<tr><td></td><td></td><td colspan="2">Margin (M)</td></tr>
<tr><td></td><td></td><td>With Margin</td><td>Without Margin</td></tr>
<tr><td rowspan="2">Orientation (O)</td><td>Portrait</td><td>Condition A

</td><td>Condition B

</td></tr>
<tr><td>Landscape</td><td>Condition C

</td><td>Condition D

</td></tr>
</table>

4.7. Procedure

The participants were briefed on the purpose of the study and were asked to read and sign the Informed Consent Form (Appendix C). They were then given written instructions on the experiments. Prior to the experiments, a screening test (Appendix A) was given to all participants. For the screening test, participants were instructed to read a passage displayed in a paper as quickly and accurately as possible although no time constraint was given. When participants finished reading the test passage, they signaled their completion to the experimenter and the time taken was measured using a stop-watch. Participants were then asked to answer five multiple-choice comprehension questions created by ETS. They were not allowed to refer back to the document when solving the questions. If participants read the passage within 5 minutes and had more than three correct answers out of five questions in the screening test, they proceeded to the experiment. The participants, who had not read within five minutes or had less than three correct answers, could not participate in the experiment.

The participants who met the performance criteria of the screening test were asked to complete a demographic sheet and a background questionnaire (Appendix E). They were then shown the PDA with a sample text and given instructions on how to use the buttons to navigate from one screen to the next. They were allowed to practice navigation until they were comfortable with use of the buttons. Participants were then assigned to one of the four conditions following the counter balanced order explained in the experimental design.

The procedures of the experiment were identical to the screening test except that the participants were reading from the PDA instead of paper. While participants were reading, the experimenter observed all actions made by participants for further behavioral analysis, such as returning to the previous page. On completion of reading task in each experimental condition, participants signaled their end of reading, and reading rate was recorded by the experimenter using a stop-watch. After reading each passage, participants answered five comprehension questions printed on a paper. After answering the questions, the subjective questionnaire form was given to measure subjective ease-of-reading, mental workload and satisfaction. After completing tasks in all experimental conditions, the subjective questionnaire was given to obtain preference (Appendix D). At the end of the experiment, a short interview was administered to obtain information regarding their experience of reading with the PDA. All participants were compensated at $7.50 per hour.

4.8. Analysis

Data for the dependent measures (reading speed, reading comprehension, reading proficiency, perception of ease of reading, mental workload, and satisfaction) were analyzed using a two-factor analysis of variance (ANOVA) with a significant level of 0.05 ($\alpha = 0.05$) to test the effect of orientation and margin, and interaction between the two. A Pearson Correlation Coefficient was calculated to investigate a possible relationship between reading time and comprehension. As a post-hoc analysis, the Bonferroni pairwise comparisons were conducted, adjusted for the family of significant differences at the .05 level. A chi-square test of independence was performed for the preference of the screen orientation, margin, and the combination of the two.

CHAPTER 5. RESULTS

The descriptive statistics and ANOVA summary table for the results of each dependent measures are presented in Appendix F and G respectively.

5.1. Reading Speed

A two-way ANOVA revealed no main effects of orientation and margin (p=0.117 for orientation and p=0.816 for margin). No significant interaction between the two factors was found (p=0.812). A detailed summary of results is shown in Appendix G. While the main effect of the orientation was not significant at $p<0.05$, the trend of data showed a greater difference of reading speed between the levels of orientation, compared to the levels of margin (Figure 4 and 5). Also, a slight difference among the four different treatment conditions was found (Figure 6).

5.2. Reading Comprehension

The comprehension measure was recorded as the percent of correct answers first, and the percent was converted using the arcsine transformation prior to the ANOVA test. Howell (1992) stated that when an experiment deals with a proportional dependent variable, the data tends to have binomial distribution where the variance is dependent on the mean. Therefore, he recommended using the arcsine transformation (Y=2 arcsine $\sqrt{p}$, where p is the proportional data) to compress the tail of the distribution. A two-way ANOVA test, which was performed with transformed data, revealed that no significant main effect of the two factors was found on reading comprehension (p=0.895 for orientation, p=0.154 for margin and p=0.300 for interaction between the two).

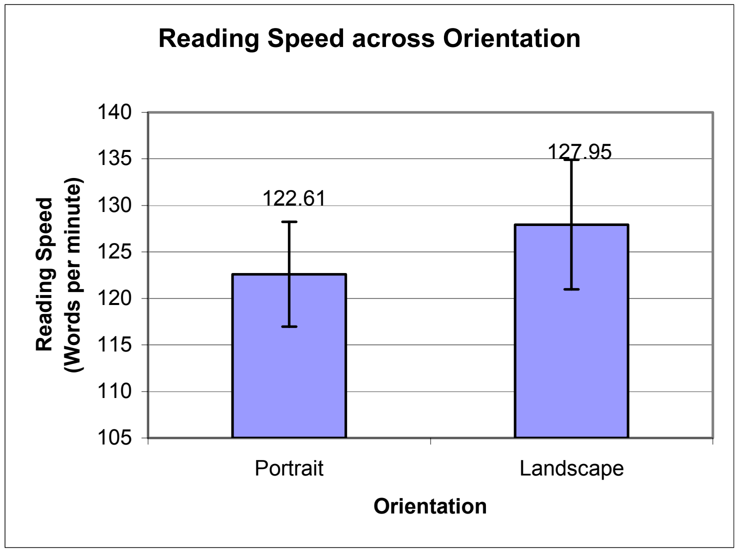


Figure 4. Mean Reading Speed across Orientations

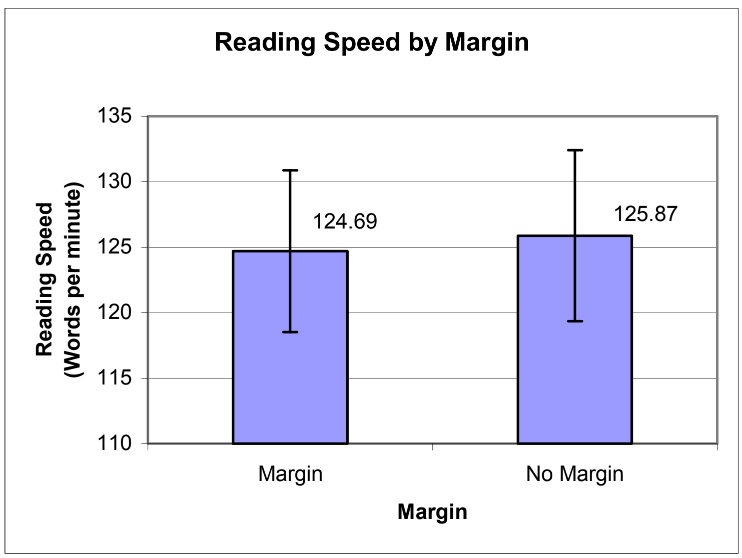


Figure 5. Mean Reading Speed by Margin

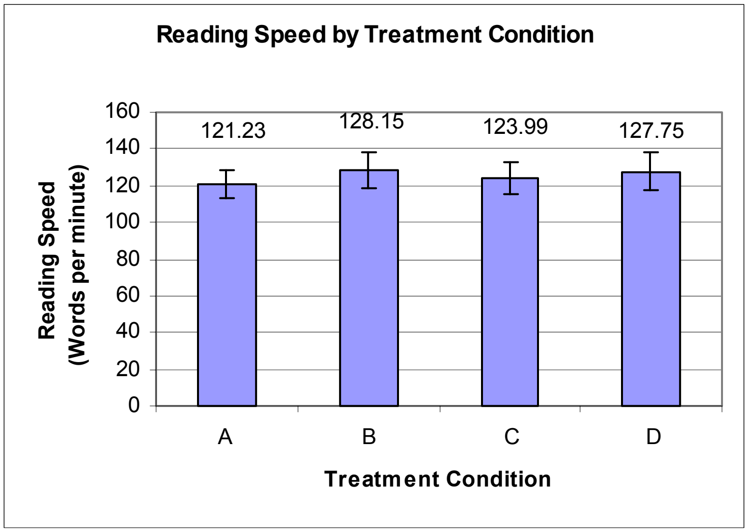


Figure 6. Mean Reading Speed by Treatment Condition

5.3. Reading Proficiency

Reading proficiency was calculated by multiplying reading speed by reading comprehension. A two-way ANOVA indicated that there was no main effect and no interaction of the two factors on reading proficiency (p=0.310 for orientation, p=0.635 for margin and p=0.659 for interaction between the two). While not significant, increased proficiency level was found from landscape over portrait mode, as shown Figure 7.

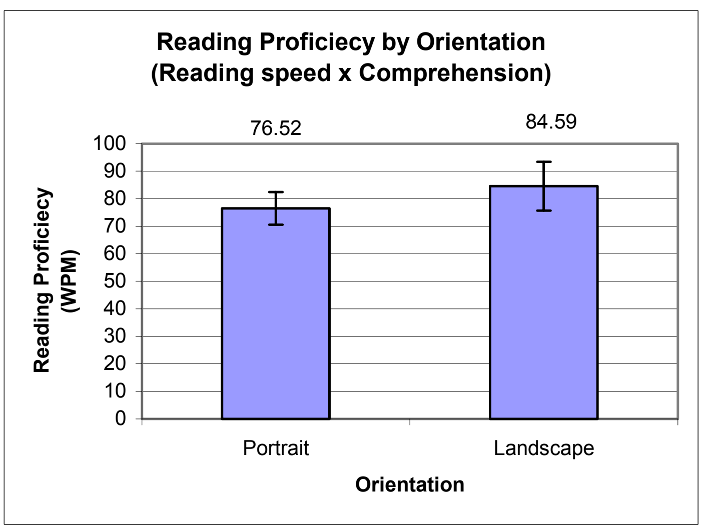


Figure 7. Reading Proficiency by Screen Orientation

Pearson's Product-Moment Correlation was calculated between reading speed and comprehension, and it did not show a significant correlation between the two measures (r=0.086, p=0.501). See Figure 8 for the scatter plot.

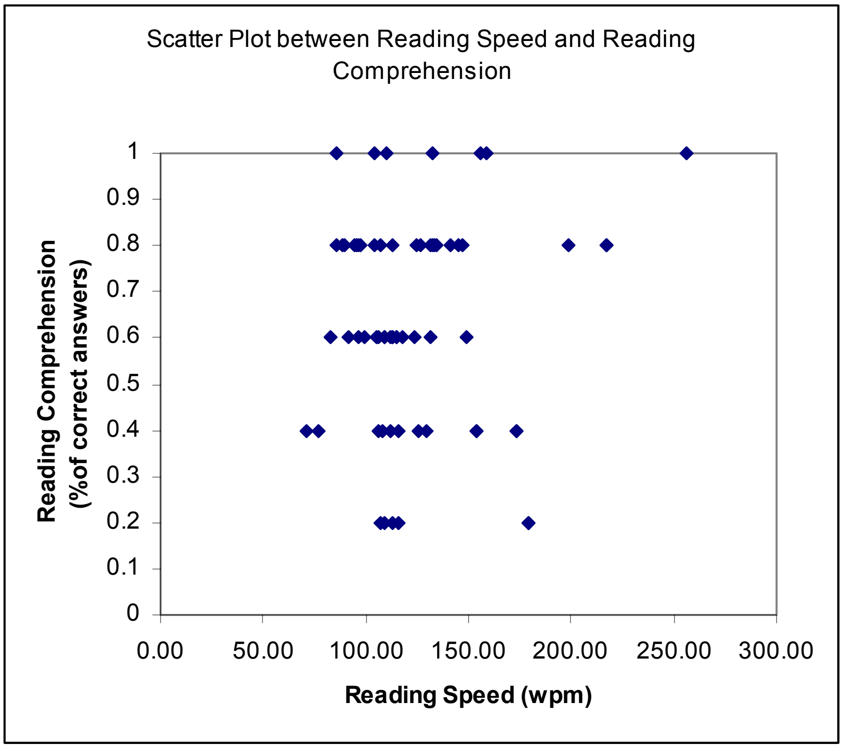


Figure 8. Scatter plot between Reading Speed and Reading Comprehension

5.4. Perception of Ease of Reading

A two-way ANOVA revealed that no significant main effect and no interaction between the two factors were found on subjective perception of ease of reading (p=0.723 for orientation, p=0.447 for margin, and p=0.361 for interaction between the two). Although the difference was not significant at $p<0.05$, the trend of data showed that the 'with margin' mode increased subjective ease of reading slightly (See Figure 9).

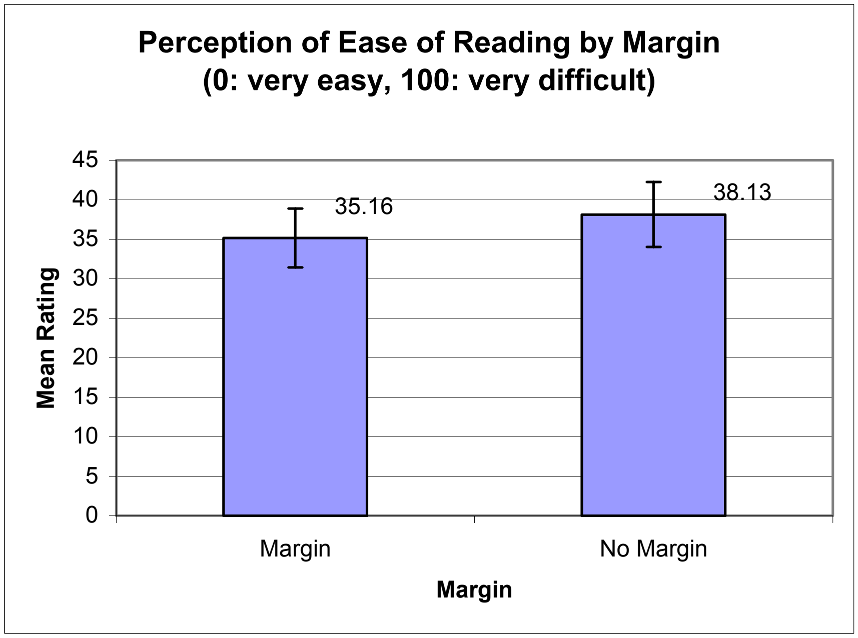


Figure 9. Subjective Ease of Reading by Margin

5.5. Mental Workload

Mental Demand

A two-way ANOVA revealed no significant main effect of orientation (p=0.606) and margin (p=0.869) as well as no interaction between the two factors (p=0.190) was found on mental.

Performance

A two-way ANOVA revealed no significant main effect of orientation and margin as well as no interaction between the two factors was found on performance (p=0.424 for orientation, p=0.114 for margin, and p=0.120 for interaction between the two).

Effort

A two-way ANOVA revealed no significant main effect of orientation and margin as well as no interaction between the two factors was found on effort (p=0.968 for orientation, p=0.939 for margin, and p=0.338 for interaction between the two).

Frustration

A two-way ANOVA revealed no significant main effect of orientation and margin as well as no interaction between the two factors was found on frustration (p=0.361 for orientation, p=0.690 for margin, and p=0.148 for interaction between the two).

Overall Mental Workload

Overall mental workload was calculated by averaging the four dimensions with the same weight. A two-way ANOVA revealed no significant main effect of orientation and margin as well as no interaction between the two factors was found on overall mental workload (p=0.506 for orientation, p=0.765 for margin, and p=0.101 for interaction between the two).

5.6. Satisfaction

A two-way ANOVA revealed no significant main effect of orientation and margin as well as no interaction between the two factors was found on satisfaction (p=0.815 for orientation, p=0.635 for margin, and p=0.273 for interaction between the two).

5.7. Preference

A Chi-square test of independence was performed to test preference of orientation, margin, and treatment conditions. Participants did not have an orientation preference as shown in Figure 10 (χ^2 = 0.00, df=1, p =1.000). Although no significant preference of margin (χ^2 = 1.00, df=1, p=0.3173) and treatment condition (χ^2 = 1.50, df=3, p=0.6823) was found in the statistical test, there was a slight difference of preference for margin and treatment conditions as shown in Figure 11 and 12 respectively.

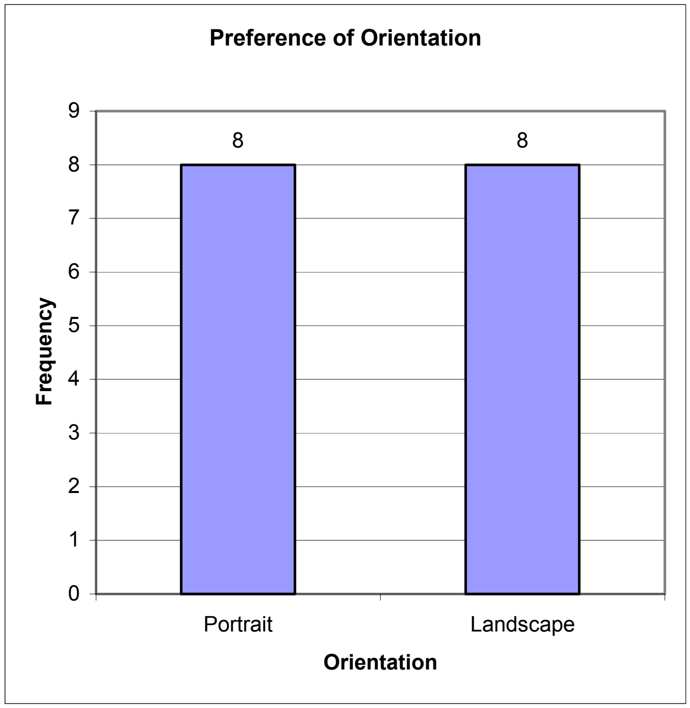


Figure 10. Preference of Orientation

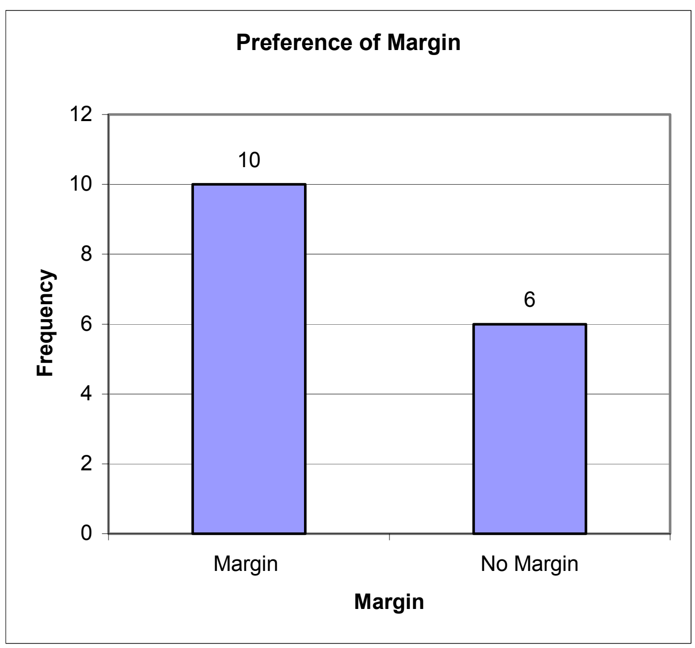


Figure 11. Preference of Margin

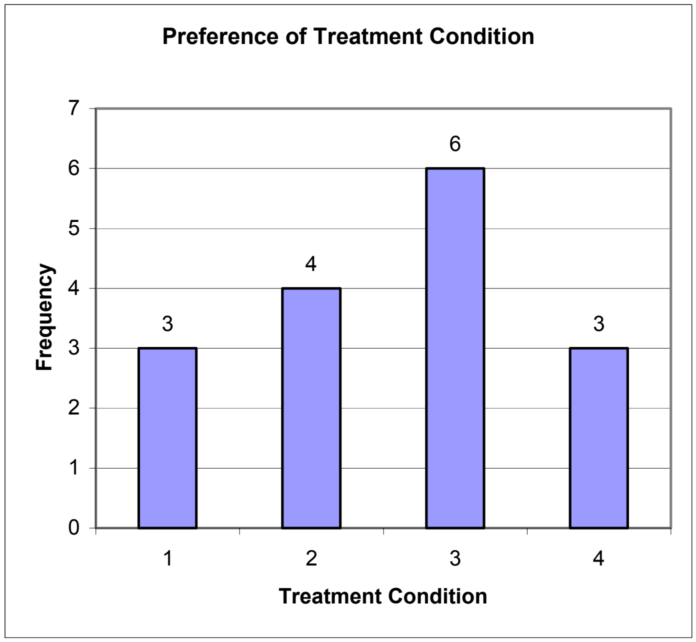


Figure 12. Preference of Treatment Condition

5.8. Supplemental ANOVA for passage effect

A supplemental ANOVA was performed to test if the content of passages influenced objective and subjective measures. A one-way ANOVA indicated that there were no significant effects of passages on all dependent measures. A statistical summary is shown in Appendix G.

CHAPTER 6. DISCUSSION

This study aimed to determine whether screen orientation and margin influence readability of text from a small hand held device screen. It was hypothesized that main effects and interaction effects of orientation and margin would be present in the objective measures that include reading speed, comprehension, and reading proficiency and subjective measures that include perception of ease of reading, mental workload, and satisfaction.

The results of this experiment indicated that no main effect of orientation and margin was found in all dependent measures. This was an unexpected result that all hypotheses were rejected. Nonetheless, several findings worth being discussed were observed from the trends in the data. These trends will be explained on the basis of each dependent measure.

6.1. Objective Measures

Reading speed

No significant effect of orientation and margin was found on reading speed. In other words, when people read text from a small hand held screen (3.8-inch screen diagonal), orientation and margin do not affect reading speed significantly. However, the trend in the data indicated that orientation affected reading speed more than margin did. It was shown that participants read faster from landscape mode compared to portrait mode (See Figure 4). This trend is consistent with the findings from previous studies, supporting the concept that longer line length produces faster reading speed. Changing orientation from portrait mode to landscape mode in this experiment increased the line length by six characters per line (cpl), getting close to the suggested optimum line length 55cpl. Indeed, increasing reading speed was found in this experiment as the line length increased, but it could not reach a significant level due to a high level of individual variability in reading speed (M=125.28, SD=35.67, MIN=70.88, MAX=256.57).

On the other hand, the effect of margin on reading speed could not be found even in the raw data. This result was also consistent with previous studies that have failed to find an effect of margin on reading speed. Therefore, it seems that margin does not promote greater reading speed in a handheld small screen.

Reading Comprehension

No significant effect of orientation and margin was found on reading comprehension. This was not surprising because it was anticipated that participants would spend more time to retain a certain comprehension level since there was no time constraint on reading passages.

Reading Proficiency

No significant effect of orientation and margin was found on reading proficiency. This measure was calculated as an overall reading performance index concerning reading speed and comprehension. Pearson's Product-Moment Correlation was calculated to investigate a possible relationship between speed and comprehension, and it showed a low level of correlation between the two measures (r=0.086, p=0.501). This implied that participants did spend more time to retain a level of comprehension.

This was not a surprising result for the same reason explained in reading comprehension. A slightly higher reading speed but no comprehension difference was found in the landscape mode, this indicates that the proficiency level in the landscape mode was also slightly higher than the portrait mode, affected only by the reading speed (Figure 7).

6.2. Subjective Measures

The results of this experiment indicated that margin and orientation do not affect subjective measures including ease of reading, mental workload, and satisfaction. These were unexpected results, since it was anticipated that subjective measures would be more sensitive than objective measures in this experiment. Especially, margin effect was expected to be significant in the case of subjective measures since past literature supported the finding that the existence of a margin would make reading easier by helping the eyes to make an accurate return sweep with increased visual contrast. Consequently, the margin was expected to decrease mental workload. Although the effect of margin was not significant in this experiment, such a trend was found in the perception of ease of reading (Figure 9). However, there was no such trend in the overall mental workload and satisfaction. Therefore, it is interpreted that a change on orientation and margin does not influence legibility and readability of text in a small hand held device screen.

One possible explanation for the lack of hypothesis confirmation about the margin effect is that it is possible that the bezel (hardware exterior case surrounding the display) might play a role as a width of margin (Figure 13). This speculation is based on the post-experiment interview with participants. Four participants mentioned that it was easy for them to read text from the screen without margin because of the gray frame of the exterior case. Therefore, it is highly possible that the gray bezel produced a certain level of visual contrast that supported the eye return to the next line. This speculation needs to be studied with another factor, the color of the case, since it is questionable how it would have worked had the case been a darker color.

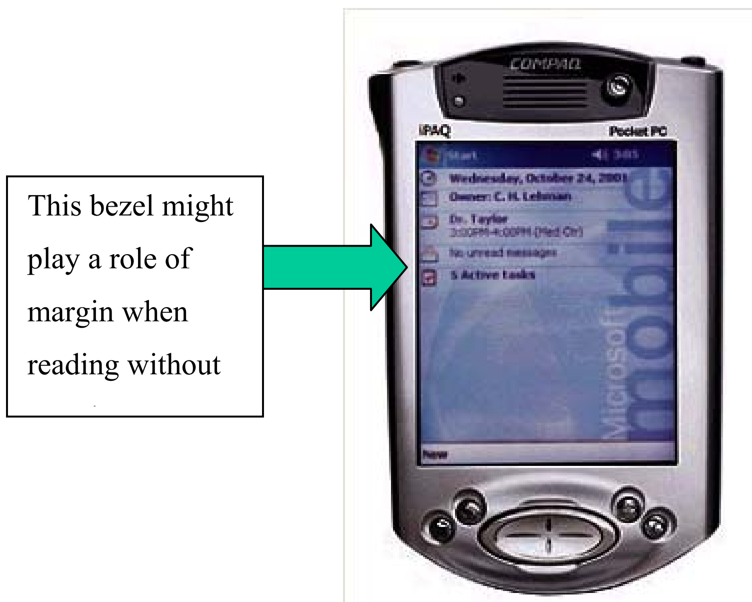


Figure 13. Hardware Exterior Case (bezel) that was supposed to play a role of margin

Another possible explanation for the lack of hypothesis confirmation about the screen orientation is that, whether participants read from portrait or landscape mode of the small screen, the passages were still reasonably readable and comprehensible. The difference in levels of orientation was too small to produce a difference with not only the objective measures but also the subjective measures. In other words, the six-character difference between the two levels of orientation was not perceived sufficiently enough to make participants feel different. The evenly distributed result for preference of orientation (50% for portrait and 50% for landscape) strongly supports this speculation.

It was an interesting finding since it contrasts with a well-known public preference for the portrait mode. Wearden (1998) reported in his study of about 200 college students that participants had clear preference for the portrait mode regardless of whether the texts were presented on paper or on screen. But the result of this study showed that no prevailing preference of screen orientation was present in a handheld device small screen, on the contrary, the landscape mode was preferred over the portrait mode when it was combined with the margin effect. This was supported by the result of preference for treatment condition although its difference failed to reach the significance level, $p<0.05$ (See Figure 12). This finding might suggest that as long as it is within a reasonable range, screen ratio is not the main concern when a handheld device with a small screen is designed for the purpose of reading. Also, it is recommended that the manufacturer of a handheld device should make it possible to rotate the screen orientation so that users can read text in their preferred orientation.

Another interesting finding was observed from the four participants who chose portrait in the post-experimental interview. They mentioned that they used to prefer landscape for computer screen and portrait for paper, but the handheld computer gave them a feeling that it would be like paper rather than a computer because they could hold it with their hands when they read. This might imply that reading from a handheld computer screen is not analogous with reading from stationary computer screen. Therefore, caution should be exercised while applying findings from a stationary computer screen to a handheld device small screen when it comes to readability of text for comprehension reading. Future comprehensive investigation should check the difference between the stationary computer screen and handheld computer small screen.

6.3. Post-Experiment Questionnaire

6.3.1. Reading Experience with The Handheld computer

As explained in the procedure section, a post-experiment questionnaire and interview were administered to obtain information regarding participants' experience of reading with the handheld computer. Overall, the results of the post-experiment questionnaire indicated that participants had a favorable experience with the handheld computer (PDA).

In responding to the question, "Did you feel frustration during experiments because of the small screen size?", 14 participants (87.5%) out of 16 answered that they did not feel any frustration caused by the small screen through experiments ($\chi^2 = 9.0$, $df=1$, $p=0.0027$). The result is shown in Figure 14. Also, in responding to the question, "Do you think a PDA is useful for reading texts?", 15 participants (95%) answered that that the PDA is useful for reading texts. The result is shown in Figure 15.

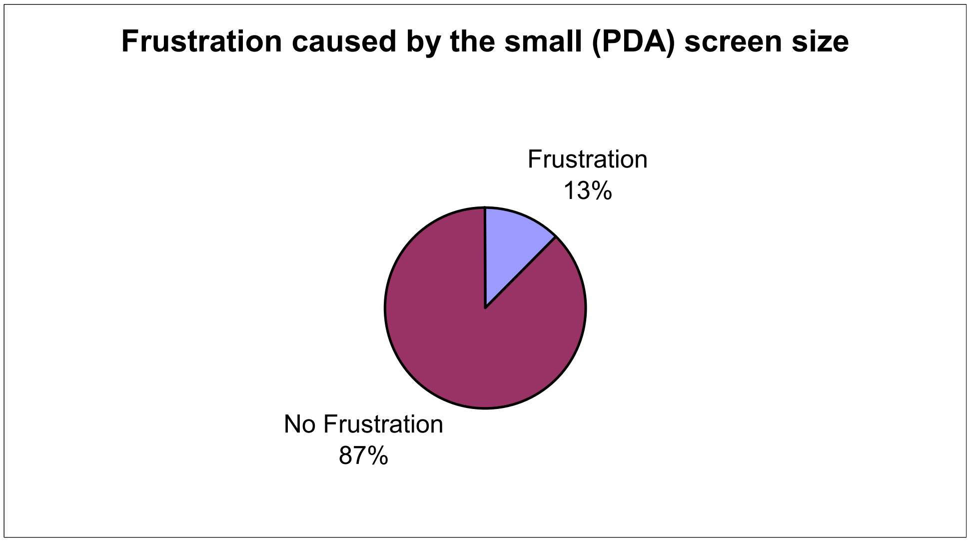


Figure 14. Percent of Participant Who Did Not Experience Frustration

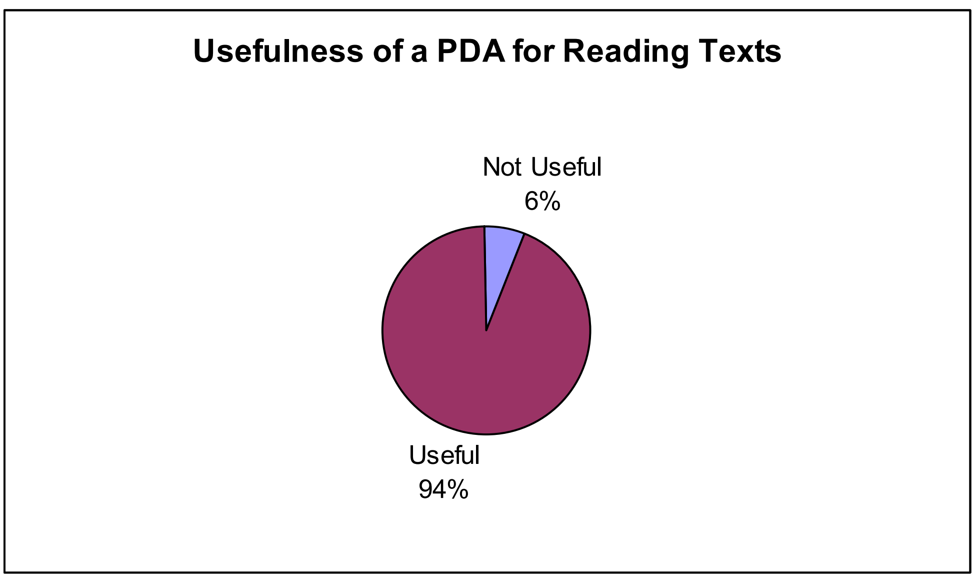


Figure 15. Usefulness of a PDA for Reading Texts

One interesting finding was observed in terms of participants' change of perception about the small screen of the PDA. To assess how participants perceived the screen size of the PDA, the same question, "Do you think the PDA shown here is too small to read text?", was asked to each participant before and after experiment. While six participants (44%) thought that the screen was too small before experiments, the number decreased to only two (13%) after experiments, as shown in Figure 16. A Chi-square test conducted with the data collected after experiments, and it showed that participants perceived the PDA screen was not too small to read a passage (χ^2= 9.0, df=1, p=0.0027). Although McNemar's Test showed that this difference between the two results (before and after experiment) was not significant, detailed data observation revealed a remarkable finding. All of the six participants, who answered, "Yes- the screen is too small" prior to the experiment, changed their opinion after the experiments. Demographic information revealed that they were all non-PDA users. This result might imply that non-PDA users tend to perceive a handheld computer screen to be too small because of the physical size, but their perception can change by having them experience the products. This might suggest that maximizing potential customers' exposure to the product is an important factor to their sale in the handheld computer market.

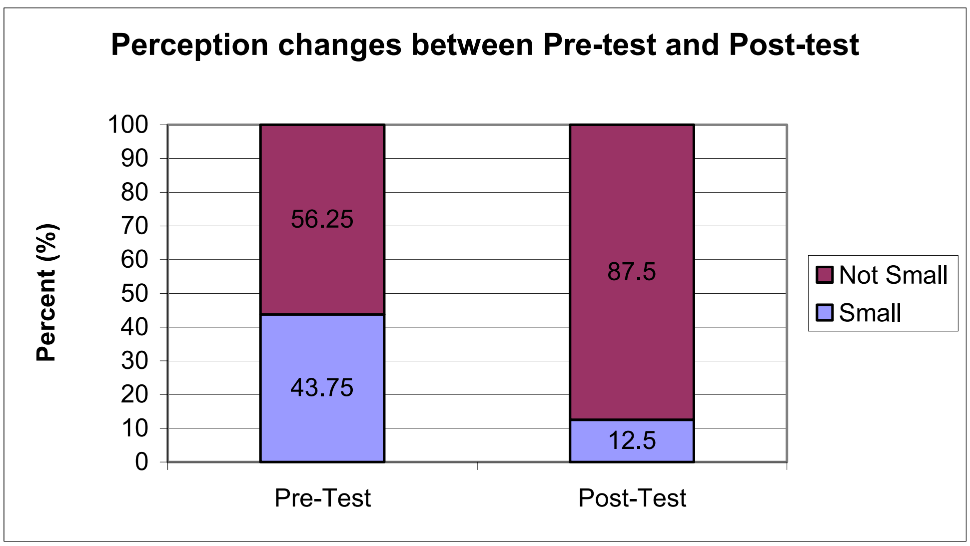


Figure 16. Perception Change on The PDA Screen Size

6.3.2. Artifactual Data

The post-experiment questionnaire also contained questions to elicit participants' views on the following issues:

1) Reading material that they would wish to read using handheld computers
2) Positive and negative comments with respect to reading on handheld computers
3) Suggestions to improve reading with handheld computers

Participants were asked to mark all genres of text they would wish to read using a handheld computer. The genre classification used in this question was proposed by Borchers (1999). The results are summarized in Table 7.

Table 7. Genres That Participants Wish to Read Using a Handheld Computer

Genre Classification	Number of Participants
Reference and documentation (i.e. dictionary, phone book, reference manual)	14
Learning material (i.e. tutorial, school books)	3
Browsing material (i.e. e-journal, newspaper, magazine)	11
Entertainment (i.e. novels, comics)	6

Table 8. Positive and Negative Comments on Reading from A Small Screen

Positive	Negative
• Much more “contained” – I felt that I was in control of reading. • Focus and concentration were greater than reading from laptop computers because of less distracting factors (i.e. stiff posture, keyboard) • It gave me kind of personal feel of PDA. • Reading is not hard and allows me to orient myself the way I want to • I can hold it at any distance (you can hold it closer to your face to compensate for its small size)	• When a paragraph is too long, continuing on to the next page, it was hard to read. • Really long documents may cause problems • Having to turn to the next page too often • Images would not work well on it • I can’t usually see the beginning and ending of a paragraph • Difficult to understand structure of passage because paragraph does not fit on one page • Color of letters is not solid like printed text • I was lost in context a bit at the beginning

They were also asked to provide positive and negative comments about reading from a handheld computer as well as suggestions to improve reading from a small screen. Positive and negative comments along with improvement suggestions for improvement are summarized in Table 8 and 9 respectively.

Table 9. Suggestions to Improve Text Display on A Small Screen

Suggestion	No. of Response
Scrolling facility (i.e. jog wheel): it increases flow of test and easy to skip multiple steps forward and backward	4
Reduce paragraph length	3
Greater overlap between pages (i.e. last 2-3 lines of previous page are the top 2-3 lines of following page)	3
Changing orientation should be optional	3
Margin width should be changeable	2
Ability to change font type and size	2
Ability to provide audio output	2

6.3.3. Design Recommendations

Based on results of the experiment and post-experiment questionnaire, a set of recommendations was developed to improve reading from handheld computers with small screens as follows:

1) Reading software should be designed to be customizable in terms of margin width, screen orientation, allowance of overlap between pages, font type and font size.

- Source: Experiment Result, Post-questionnaire; No difference of reading performance and preference was found.

2) Both text presentation modes (page and scrolling mode) should be provided.

- Source: Post-questionnaire data: Despite clear preference of page mode, Scrolling increases text flow and is easy to skip multiple steps forward and backward

3) Make the length of paragraph shorter if possible.

- Source: Post-questionnaire data: Difficult to understand the structure of a passage because a paragraph does not fit on one page

CHAPTER 7. LIMITATIONS

Although this study attempted to determine the effect of screen orientation and margin that may have contributed towards an improved reading performance and satisfaction when people are reading text from a handheld device with a small screen, some issues may limit the outcome of this study. First, there are many variants of handheld computers available in the market, called palmtops, handheld computers, communicators, etc. Designed to fit in one's pocket, all of the devices are smaller than a paperback book, but each of the devices differs in functionality and screen characteristics, such as display type, screen size and resolution. This study used only one such PDA model being sold in the market, the Compaq iPAQ 3850. Therefore, generalization of the results may be limited to product-specific features, such as the screen size, screen aspect ratio, display type and screen resolution, etc. With rapid advances in technology, some cell phone products also qualify as handheld computers, but regular cell phone display sizes displaying about three lines of words will be excluded in the definition of hand held computers in the case of this study. It must be reiterated that the handheld device small screen in this study was a 3.8-inch screen (8.17cm x 6.13cm) of a Compaq iPAQ 3850(PDA) with a resolution of 320 x 240 pixels.

Second, it must be emphasized that the task used in this study involved a detailed reading of academic articles. As introduced in the literature review section, there were five different reading types (See Table 1). The readings used in this experiment can be classified as the memorizing reading type that necessitates remembering the idea and thoughts of a passage. Many inquiries of the GRE test used in this experiment questioned about the authors' implications in the corresponding passages. In addition, the manner in which the participants were not allowed to refer back to the passage made participants try to memorize certain ideas and names in the passages. Although it was expected that participants would perform a normal reading (rauding) by providing an instruction ("read as quickly and accurately as possible"), the mean of reading speed obtained from participants, 125.28 wpm, indicated that they actually performed the memorizing reading type (See Table 3).

It is obvious that readers use a different reading type based on their goals. Scanning and skimming reading type are totally different in that it requires very quick saccadic movement, compared to rauding, learning and memorizing reading types. Therefore, the results of this study cannot be applied to scanning or skimming tasks with handheld computers.

CHAPTER 8. FUTURE RESEARCH

8.1. Further Studies with Increased Statistical Power

Although this study found no significant effect of orientation and margin in any dependent measure and preference, the small number of participants used in this study raises issues in terms of statistical power. For example, the data trend in reading speed indicated with a low p-value (0.117) that landscape tends to produce faster reading than portrait does. It might imply that a main effect of screen orientation in reading speed could be found in a study with increased number of participants. Also, preference of orientation and margin as well as subjective impression such as frustration, usefulness and perception of the screen size were obtained from only 16 participants. These issues must be addressed using a large number of participants to increase statistical power.

8.2. Sequential Experimentation with More Factors and Reading Type

Since the early 1980s', there have been a number of empirical studies investigating the factors that affect text readability on a computer screen. These studies identified a number of factors such as line length, window height, font size, margin width, interline spacing, font type, anti-aliasing, resolution, display type, color, text manipulation, etc. Nevertheless, all of the previous studies have employed a traditional experimental design concerning two or three factors as in the present study. Consequently, a large number of significant independent variables have been reported, but the functional relationships among the variables for objective and subjective reading experiences remain unknown.

As described in the literature review, reading is a highly complex perceptual and cognitive process that is influenced by many factors. Therefore a comprehensive study should be conducted to take into account as many factors as possible to investigate the optimum computer screen design for reading with computers, regardless of the screen size. According to my review, few studies have been conducted so far with not only a handheld computer screen but also a desktop computer screen, and in here lays a future research opportunity.

As a large-scale factorial design faces time and cost limitations in handling a large data set, the sequential experimentation procedure is appropriate for a reading study. See Williges, Williges, & Han (1992) for detailed information on sequential experimentation. This sequential experimentation should be applied to different reading tasks such as scanning reading type for a handheld computer user surfing the Internet in a mobile environment. It also can be applied to the rauding or memorizing reading type for academic use in education. At the end of sequential experimentation, we can obtain not only an empirical model of each reading type but also quantitative guidelines for designing reading user interfaces.

8.3. Usability Issues related to Reading from Handheld Computers

Usability of handheld computers in reading text cannot be overlooked. Reading texts from a handheld computer requires many different activities from the users, such as downloading texts, navigation through menus, using other functionalities (highlighting or taking memo, etc.). As provided in the artifactual data section, various issues related to improving reading from a small screen have already been raised by a small number of people. Few studies have been conducted on this usability issue until now. Consequently, a set of usability dimensions to evaluate usability of handheld computers in educational environment has not yet been defined. This is another study that should be conducted in the future.

CHAPTER 9. CONCLUSION

This study was conducted to investigate the main effect of orientation and margin of a handheld computer small screen on a detailed reading task performance and preference. The result of this study can be summarized as follows. First, neither screen orientation nor margin had a significant influence on objective reading performance, such as reading speed and reading comprehension. Second, neither screen orientation nor margin affects subjective measure, such as perception of ease of reading, mental workload, and satisfaction. Third, there was no prevailing preference of screen orientation and margin in reading texts from a handheld computer screen.

In spite of the lack of hypothesis confirmation throughout the experiment, this study identified several issues. First, screen orientation and margin are not the main issue in designing reading software on a handheld computer for detailed reading task. Since there is no clear preference on the both factors, it is recommended that a document viewer on a handheld computer should be designed to be customizable so that users can read text in their preferred mode. Second, screen ratio is not a main issue in hardware design of handheld reading device as well. For reading purpose, there was no effect of screen ratio (3:2 in landscape mode and 2:3 in portrait mode). Third, reading from a handheld computer screen is different than reading from a stationary computer screen in terms of effects of the independent factors on reading performance and preference. Therefore, findings from a stationary computer screen should be used with caution when designing a handheld device itself or reading software for a small screen. Above all, a comprehensive study dedicated to text readability and usability of a handheld computer should be conducted simultaneously for people to facilitate their information access more efficiently using their handheld computers.

REFERENCES

Bernard, M., Fernandez, M., & Hull, S. (2002). The Effects of Line Length on Children and Adults' Online Reading Performance. *Usability News*.

Bernard, M., & Mills, M. (2002). So, What size and type of font should I use on my website? *Usability News*.

Borchers, J. O. (1999). Electronic Books: Definition, Genres, and Interaction Design Patterns. In *ACM CHI 1999, Electronic Book Workshop*. Pittsburgh, PA.

Boyarski, D., Neuwirth, C., Forlizzi, J., & Regli, S. H. (1998). A study of fonts designed for screen display. In *CHI 98 Conference Proceedings* (pp. 87-94).

Buchanan, G., Farrant, S., Jones, M., Marsden, G., Pazzani, M., & Thimbleby, H. (2001). Improving Mobile Internet Usability. In *In Proceedings of the Tenth International World Wide Web Conference* (pp. 673-680). Hong Kong.

Carver, R. P. (1990). *Reading Rate. A review of Research and Theory*. New York: Academic Press, Inc.

De Bruijn, D., De Mul, S., & Van Oostendorp, H. (1992). The influence of screen size and text layout on the study of text. *Behaviour and Information Technology*, *11*(2), 71-78.

Deneen, L. (2001). Handheld PDAs and Wearable Computing Devices. *EDUCAUSE*.

Dillon, A. (1992). Reading from paper versus screens: a critical review of the empirical literature. *Ergonomics*, *35*, 1297-1326.

Dillon, A., Richardson, J., & McKnight, C. (1990). The effect of display size and text splitting on reading lengthy text from the screen. *Behaviour and Information Technology*, *9*(3), 215-227.

Duchnicky, R. L., & Kolers, P. A. (1983). Readability of text scrolled on visual display terminals as a function of window size. *Human Factors*, *25*(6), 683-692.

Dyson, M. C., & Kipping, G. J. (1998). The effects of line length and method of movement on patterns of reading from screen. *Visible Language*, *32*, 150-181.

Egan, D., Remde, J., Landauer, T., Lochbaun, C., & Gomez, L. (1989). Behavioural evaluation and analysis of a hypertext browser. In *Proceedings of CHI* (pp. 205-210). '89, New York: Association of Computing Machinery.

Elkerton, J., & Williges, R. C. (1984). Information retrieval strategies in a file search environment. *Human Factors*, *26*, 171-184.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: The MIT Press.

ETS. (2002). *GRE Practicing to take the General Test*. (10 ed.): Educational Testing Services.

Felici, J. (2000). Clear Type, CoolType: The Eyes Have It. *The Seybold Report (Internet Publishing)*, *4*, http://www.seyboldreports.com/SRIP/free/0408/cooltype.html.

Fung, P., Hennessy, S., & O'Shea, T. (1998). Pocketbook Computing: A Paradigm Shift? *Computers in the Schools*, *14*(3/4), 109-118.

Gould, J. D., Alfaro, L., Finn, R., Haupt, B., & Minuto, A. (1987). Reading from CRT displays can be as fast as reading from paper. *Human Factors*, *29*, 497-517.

Hansen, W. J., & Haas, C. (1988). Reading and writing with computers: a framework for explaning defferences in performance. *Communications of the ACM*, *31*, 1080-1089.

Hart, S. G., & Staveland, L. E. (Eds.). (1988). *Development of Nasa-TLX: Results of empirical and theoretical research.* V.V.: North-Holland: Elsevier Science Publishers.

Hennessy, S. (2000). Graphing investigations using portable (Palmtop) technology. *Journal of Computer Assisted Learning*, *16*, 243-258.

Howell, D. C. (1992). *Statistical methods for psychology*. (3rd ed.). Boston: Duxbury Press.

Jones, M., Marsden, G., Mohd-Nasir, N., & Boone, K. (1999). Improving Web Interaction on Small Displays. In *proceedings of 8th International World-Wide Web Conference* (pp. 51-59).

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329-354.

Kak, A. V. (1981). Relationships between readability of printed and CRT-displayed text. In *Proceedings of Human Factors Society - 25th Annual Meeting* (pp. 137-140). Santa Monica: CA:Human Factors Society.

Kang, T. J., & Muter, P. (1989). Reading Dynamically Displayed Text. *Behaviour & Information Technology*, *8*(1), 33-42.

Kruk, R. S., & Muter, P. (1984). Reading of continuous text on video screens. *Human Factors*, *26*, 339-345.

Legge, G. E., G., P. D., Rubin, G. S., & Schleske, M. M. (1985). Psychophysics of reading-I. Normal vision. *Vision Research*, *25*(2), 239-252.

Lindsay, P. H., & Norman, D. A. (1972). *Human information processing; an introduction to psychology*. New York: Academic Press.

McKnight, C., Dillon, A., Richardson, J., Haraldsson, H., & Spinks, R. (Eds.). (1992). *Information access in different media: an experimental comparison*.

Mills, C. B., & Weldon, L. J. (1987). Reading text from computer screens. *ACM Computing Surverys*, *50*, 597-624.

Morris, M., & Hinrichs, R. (1996). *Web Page Design*. California: Sun Microsystems Inc.

Muter, P. (Ed.). (1996). *Interface Design and Optimization of Reading of Continuous Text*. Norwood, N.J.: Ablex.

Muter, P., Latremouille, S. A., Treurniet, W. C., & Beam, P. (1982). Extended reading of continous text on television screens. *Human Factors*, *24*, 501-508.

Muter, P., & Maurutto, P. (1991). Reading and skimming from computer screens and books: The paperless office revisited? *Behaviour & Information Technology*, *10*, 257-266.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.

O'Hara, K., & Sellen, A. J. (1997). A Comparison of Reading Paper and On-Line Documents. In *CHI '97, Conference on Human Factors in Computing Systems*. Atlanta, GA.: ACM.

Olshavsky, J. (1977). Reading as problem solving: an investigation of strategies. *Reading Research Quarterly*, *4*, 654-674.

Oquist, G. (2001). *Adaptive Rapid Serial Visual Presentation.*Unpublished Thesis, Uppsala University, Uppsala, Sweden.

Paterson, D. G., & Tinker, M. A. (1940). *How to Make Type Readable*: Harpers.

Piolat, A., Roussey, J., & Thunin, O. (1997). Effects of screen presentation on text reading and revising. *International Journal of Human-Computer Studies*, *47*, 565-589.

Rahman, T., & Muter, P. (1999). Designing an interface to optimize reading with small display windows. *Human Factors*, *1*(1), 106-117.

Rayner, K., & Pollatsek, A. (1989). *The psychology of Reading*. Englewood Cliffs, NJ: Prentice Hall.

Reisel, J. F., & Shneiderman, B. (1987). *Is bigger better? The effects of display size on program reading*. Amsterdam: Elsevier Science Publishers B.V.

Reynolds, L. (1984). The legibility of printed scientific and technical information. In R.E.H. Zwaga (Ed.), *Information Design* (pp. 187-208). New York: John Wiley and Sons Ltd.

Robeck, M. C., & Wallace, R. R. (1990). *Psychology of reading: An interdisciplinary approach.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, NJ: Erlbaum.

Schwarz, E., Beldie, I. P., & Pastoor, S. (1983). A Comparison of paging and scrolling for changing screen contents by inexperienced users. *Human Factors*, *25*(3), 279-282.

Shepard, W. (2002). Supporting and Integrating PDA Technology on Campus. *EDUCAUSE QUARTERLY*, *3*, 67-69.

Siegel, D. (1997). *Creating Killer Web Sites*. (2nd edition ed.): United States: Haden Books.

Tinker, M. A. (1963). *Legibility of print*. Ames, IA: University of Iowa Press.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Warren, M. (2002). Why is Casio E-200 Screen is Crispy Clear.

Wearden, S. (1998). Landscape vs. Portrait Formats: Assessing Consumer Preferences. *Future of Print Media Journal*(Summer).

Williges, R. C., Williges, B. H., & Han, S. H. (1992). Developing quantitative guidelines using integrated data from sequential experiments. *Human Factors*, *34*, 399-408.

Youngman, M., & Scharff, L. (1998). Text length and margin length influences on readability of GUIs: Southwest Psychological Association.

Appendix A: SAMPLE READING TEST

Sample Reading Test

Directions: Each passage in this group is followed by questions based on its content. After reading a passage, choose the best answer to each question. Answer all questions following a passage on the basis of what is stated or implied in that passage.

Tocqueville, apparently, was wrong. Jacksonian America was not a fluid, egalitarian society where individual wealth and poverty were ephemeral conditions. At least so argues E. Pessen in his iconoclastic study of the very rich in the United States between 1825 and 1850.

Pessen does present a quantity of examples, together with some refreshingly intelligible statistics, to establish the existence of an inordinately wealthy class. Though active in commerce or the professions, most of the wealthy were not self-made, but had inherited family fortunes. In no sense mercurial, these great fortunes survived the financial panics that destroyed lesser ones. Indeed, in several cities the wealthiest one percent constantly increased its share until by 1850 it owned half of the community's wealth. Although these observations are true, Pessen overestimates their importance by concluding from them that the undoubted progress toward inequality in the late eighteenth century continued in the Jacksonian period and that the United States was a class-ridden, plutocratic society even before industrialization.

1. According to the passage, Pessen indicates that all of the following were trued of the very wealthy in the United States between 1825 and 1850 EXCEPT:

(A) They formed a distinct upper class.
(B) Many of them were able to increase their holdings.
(C) Some of them worked as professionals or in business.
(D) Most of them accumulated their own fortunes.
(E) Many of them retained their wealth in spite of financial upheavals.

2. The author's attitude toward Pessen's presentation of statistics can be best described as

(A) disapproving
(B) shocked
(C) suspicious
(D) amused
(E) laudatory

3. Which of the following best states the author's main point?

(A) Pessen's study has overturned the previously established view of the social and economic sturucture of early nineteenth-century America.

(B) Tocqueville's analysis of the United States in the Jacksonian era remains the definitive account of this period.

(C) Pessen's study is valuable primarily because it shows the continuity of the social system in the United States throughout the nineteenth century.

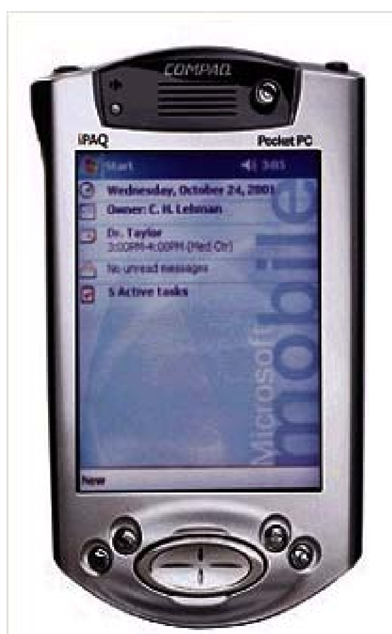
(D) The social patterns and political power of the extremely wealthy in the United States between 1825 and 1850 are well documented.

(E) Pessen challenges a view of the social and economic system in the United States from 1825 to 1850, but he draws conclusions that are incorrect.

Appendix B: A PDA Model used in this study

Compaq iPAQ 3850 series

- Operating System: Microsoft Pocket PC 2002
- Processor: 206 MHz Intel Strong ARM 32-bit RISC Processor
- Display Type: Color reflective thin film transistor (TFT) LCD, 64K colors
- Touch Screen: Yes
- Resolution: 240 x 320
- Pixel Pitch: .24 mm
- Viewable Image Size: 2.26 x 3.02 inches
- RAM: 64 MB
- ROM: 32 MB
- Input Method: Handwriting recognition, soft keyboard, voice record, inking
- Communications Port: Interface with USB / Serial connectivity that connects via serial or USB cable
- Card Slot: SD Memory Slot, Optional expansion packs
- Wireless Connectivity: Infrared port (115 Kbps)
- Speaker and Microphone: Yes
- Audio Out Jack: Yes (3.5 mm Stereo)
- Battery: 1400 mAh Lithium Polymer
- Dimensions: 5.3" x 3.3" x .62"
- Weight: 6.7 oz.



Compaq iPAQ 3850

Appendix C: INFORMED CONSENT FORM

Title of Project: Improving Text Reading on PDAs (Personal Data Assistants)
Principal Investigator: Young S. Lee
Other Investigators: Tonya L. Smith-Jackson

PURPOSE OF PROJECT

This project will investigate the effect of the small screen size of a PDA on reading performance (reading rate and comprehension) and subjective reading experience (subjective workload, ease-of-reading, satisfaction). Also, this project will examine reading software interface design to enhance reading experience with PDAs.

INFORMATION

In this project, you will participate in one of two experiments or both two experiments. Experiment 1 is a reading experiment in which you will be asked to read passages with a laptop computer and a PDA, followed by questionnaires. In Experiment 2, you will be asked to read different passages with three different reading software products for PDAs.

RISKS

Participation in this project does not place you at more than minimal risk of harm.

BENEFITS

You will be compensated for your participation, and you will be given information to contact the principal investigator to get information about the outcomes of the study. You will also benefit from knowing that you have participated in worthwhile research that has immediate and positive applications.

CONFIDENTIALITY

The information gained in this research project will be kept strictly confidential. At no time will the researchers release the results of the study to anyone other than individuals working on the project without your written consent.

You will be identified only by a 3 digit study code. Data will be stored securely and will be made available only in the context of research publications and discussion. No reference will be made in oral or written reports that could link you to the data nor will you ever be identified as a participant in the project.

COMPENSATION

You will be compensated at the rate of $7.50 per hour for participation in this research.

FREEDOM TO WITHDRAW

You are free to withdraw from this study at any time without penalty.

APPROVAL

This research project has been approved, as required, by the Institutional Review Board for Research Involving Human Subjects at Virginia Polytechnic Institute and State University and by the Department of Industrial and Systems Engineering.

PARTICIPANT'S RESPONSIBILITIES

It is very important that you keep the activities and information discussed confidential, since others will be participating in this research.

QUESTIONS

If you have questions, or do not understand information on this form, please feel free to ask them now.

PARTICIPANT'S PERMISSION

I have read and understand the Informed Consent and conditions of this project. I have had all questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project.

If I participate, I may withdraw at any time without penalty.

Signature ____

Date ____

CONTACT

If you have questions at any time about the project or the procedures, you may contact the principal investigator, Young Seok Lee at 231-9237 or youngLee@vt.edu (547 Whittemore).

If you feel you have not been treated according to the descriptions in this form, or your rights as a participant have been violated during the course of this project, you may contact Dr. David Moore, Chair of the Institutional Review Board Research Division at 231-4991.

Appendix D: Subjective Questionnaire

Subjective Preference Rating Scale and Questionnaire

Screen Orientation

1) What would you prefer?

□ Landscape   □ Portrait

Margin

1) What would you prefer? (Show screen capture picture)

□ With Margin   □ Without Margin

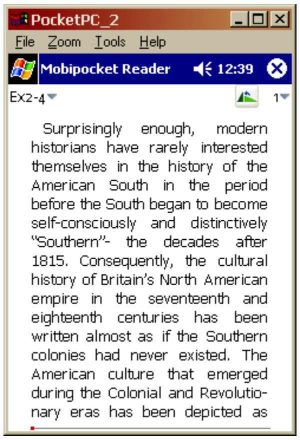
2) How helpful did you find the enough margin width when you read?

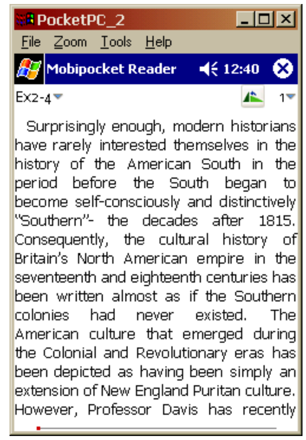
1	2	3	4	5
Very Little	Little	Moderate	Much	Very Much

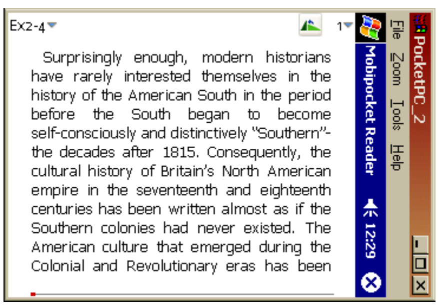
3) What would you select if you were given… (Please ask experimenter)
□ More Texts without Margin   □ Less Texts with Margin

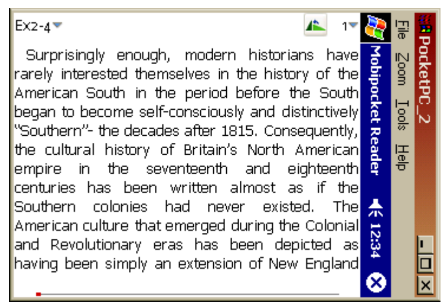
Overall (Orientation + Margin)

1) What would you prefer?







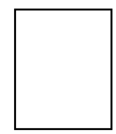


Overall Opinion

1) Overall, I think the PDA screen is too small for reading a passage.
   □ Yes   □ No

2) I was frustrated during experiment
   □ Yes   □ No

3) If you can, what changes would you make to the PDA? (based on the picture)
   a) Extend the line length (horizontal size)
   b) Extend the window height (vertical size)
   c) Screen resolution
   d) others ________________________



4) Do you think a PDA is useful for reading texts?
   □ Yes   □ No

5) If you had a PDA, what would you read using a PDA?
   a) Reference and documentation (ex: dictionary, phone book, reference manual)
   b) Learning material (ex: tutorials, school books)
   c) Browsing material (ex: e-journal, newspaper, magazines)
   d) Entertainment (ex: novels, comics)
   e) None

6) Any other comments about use of PDA for reading (reading interface design issues)?

Thanks for your participation!!

Appendix E: Background Questionnaire

Background Questionnaire

Participant Number: __________

1) Age: ________

2) Gender: _________

3) Current Grade:

- □ Junior
- □ Senior
- □ Master
- □ Ph.D

4) Do you own your PDA? Yes / No

4-1) If you own a PDA, what type of PDA is it?
□ Palm □ PocketPC □ Others _____________

4-2) How long have you used your PDA? _________years _______ months

4-3) How often do you read texts from PDA a day?
□ None □ Once □ Twice □ More than three times

4-4) What do you read from the PDA? (mark all)
□ Reference and documentation (ex: dictionary, phone book, reference manual)
□ Learning material (ex: tutorials, school books)
□ Browsing material (ex: e-journal, newspaper, magazines)
□ Entertainment (ex: novels, comics)
□ None

5) I think the PDA screen is too small to read texts
□ Yes □ No

Appendix F: Descriptive Statistics of Dependent Measures

Objective Dependent Measures

Dependent Measures	Orientation	Margin	Mean	SEMean	StDev	Variance	N
Reading Speed	Portrait	Margin	121.23	7.64	30.56	933.86	16
	Portrait	No Margin	123.99	8.54	34.15	1166.45	16
	Landscape	Margin	128.15	9.89	39.57	1565.62	16
	Landscape	No Margin	127.75	10.14	40.56	1644.86	16
Reading Comprehension	Portrait	Margin	0.69	0.06	0.24	0.06	16
	Portrait	No Margin	0.59	0.06	0.22	0.05	16
	Landscape	Margin	0.65	0.06	0.24	0.06	16
	Landscape	No Margin	0.63	0.06	0.23	0.05	16
Reading Proficiency	Portrait	Margin	80.21	8.32	33.29	1108.16	16
	Portrait	No Margin	72.84	8.65	34.6	1197.18	16
	Landscape	Margin	85.13	13.04	52.14	2718.84	16
	Landscape	No Margin	84.06	12.38	49.54	2453.97	16

Subjective Dependent Measures (1)

Dependent Measures	Orientation	Margin	Mean	SEMean	StDev	Variance	N
Perception of Ease of Reading	Portrait	Margin	33.75	5.25	21.02	441.677	16
	Portrait	No Margin	38.75	5.11	20.45	418.33	16
	Landscape	Margin	36.56	5.45	21.81	475.73	16
	Landscape	No Margin	37.5	6.60	26.39	696.67	16
Mental Demand	Portrait	Margin	35.44	5.44	21.75	473.06	16
	Portrait	No Margin	39.56	5.6	22.38	501.06	16
	Landscape	Margin	40.75	5.31	21.24	451	16
	Landscape	No Margin	37.81	6.22	24.9	619.9	16
Performance	Portrait	Margin	37.81	5.84	23.38	546.56	16
	Portrait	No Margin	39.5	6.71	26.86	721.33	16
	Landscape	Margin	44.63	6.34	25.34	642.25	16
	Landscape	No Margin	37.38	6.42	25.67	658.92	16
Effort	Portrait	Margin	40.94	5.07	20.27	410.73	16
	Portrait	No Margin	43.63	5.68	22.72	516.25	16
	Landscape	Margin	43.25	5.46	21.83	476.33	16
	Landscape	No Margin	41.06	6.2	24.78	614.06	16
Frustration	Portrait	Margin	31.44	5.58	22.31	497.73	16
	Portrait	No Margin	33.56	5.85	23.39	547.06	16
	Landscape	Margin	37.63	6.44	25.75	662.92	16
	Landscape	No Margin	32.81	6.37	25.49	649.9	16

Subjective Dependent Measures (2)

Dependent Measures	Orientation	Margin	Mean	SEMean	StDev	Variance	N
Overall Work Load	Portrait	Margin	36.41	5.01	20.02	400.81	16
	Portrait	No Margin	39.06	5.42	21.68	469.90	16
	Landscape	Margin	41.56	5.51	22.02	484.90	16
	Landscape	No Margin	37.27	5.95	23.81	566.71	16
Satisfaction	Portrait	Margin	68.75	4.82	19.28	371.67	16
	Portrait	No Margin	65.50	4.66	18.64	347.33	16
	Landscape	Margin	66.00	3.69	14.76	218.00	16
	Landscape	No Margin	66.56	4.00	15.99	255.73	16

Appendix G: Detailed Statistical Results

ANOVA Test Results for All Dependent Variables

ANOVA RESULTS:

Factor	Type	Levels	Values						
Orientat	fixed	2	0	1					
Margin	fixed	2	0	1					
Subject	random	16	1	2	3	4	5	6	7
			8	9	10	11	12	13	14
			15	16					

1) Reading Speed

Source	DF	SS	MS	F	P	
Orientat	1	455.9	455.9	2.76	0.117	
Margin	1	22.3	22.3	0.06	0.816	
Subject	15	61098.9	4073.3	**		
Orientat*Subject	15	2474.4	165.0	0.24	0.995	
Margin*Subject	15	5964.0	397.6	0.59	0.842	
Orientat*Margin	1	39.6	39.6	0.06	0.812	
Error	15	10124.6	675.0			
Total	63	80179.6				

2) Reading Comprehension

Source	DF	SS	MS	F	P	
Orientat	1	0.0037	0.0037	0.02	0.895	
Margin	1	0.5111	0.5111	2.25	0.154	
Subject	15	13.4774	0.8985	2.86	0.028	x
Orientat*Subject	15	3.0331	0.2022	1.77	0.140	
Margin*Subject	15	3.4016	0.2268	1.98	0.098	
Orientat*Margin	1	0.1315	0.1315	1.15	0.300	
Error	15	1.7148	0.1143			
Total	63	22.2732				

3) Reading Proficiency

Source	DF	SS	MS	F	P	
Orientat	1	1041.9	1041.9	1.11	0.310	
Margin	1	285.0	285.0	0.24	0.635	
Subject	15	68170.5	4544.7	3.30	0.034	x
Orientat*Subject	15	14139.8	942.7	1.21	0.358	
Margin*Subject	15	18187.6	1212.5	1.56	0.200	
Orientat*Margin	1	158.1	158.1	0.20	0.659	
Error	15	11674.4	778.3			
Total	63	113657.2				

4) Perception of Ease of Reading

Source	DF	SS	MS	F	P	
Orientat	1	9.77	9.77	0.13	0.723	
Margin	1	141.02	141.02	0.61	0.447	
Subject	15	24783.98	1652.27	7.14	0.001	x
Orientat*Subject	15	1121.48	74.77	1.01	0.496	
Margin*Subject	15	3465.23	231.02	3.11	0.018	
Orientat*Margin	1	66.02	66.02	0.89	0.361	
Error	15	1115.23	74.35			
Total	63	30702.73				

5) Mental Demand

Source	DF	SS	MS	F	P	
Orientat	1	50.8	50.8	0.28	0.606	
Margin	1	5.6	5.6	0.03	0.869	
Subject	15	23336.5	1555.8	5.61	0.001	x
Orientat*Subject	15	2746.5	183.1	1.73	0.150	
Margin*Subject	15	3004.6	200.3	1.89	0.114	
Orientat*Margin	1	199.5	199.5	1.88	0.190	
Error	15	1587.7	105.8			
Total	63	30931.2				

7) Performance (Mental Workload)

Source	DF	SS	MS	F	P	
Orientat	1	87.9	87.9	0.67	0.424	
Margin	1	123.8	123.8	2.82	0.114	
Subject	15	34160.9	2277.4	39.97	0.055	x
Orientat*Subject	15	1955.9	130.4	1.11	0.420	
Margin*Subject	15	659.0	43.9	0.37	0.967	
Orientat*Margin	1	319.5	319.5	2.72	0.120	
Error	15	1760.2	117.3			
Total	63	39067.1				

8) Effort (Mental Workload)

Source	DF	SS	MS	F	P	
Orientat	1	0.25	0.25	0.00	0.968	
Margin	1	1.00	1.00	0.01	0.939	
Subject	15	24090.94	1606.06	7.39	0.001	x
Orientat*Subject	15	2233.75	148.92	1.53	0.208	
Margin*Subject	15	2480.00	165.33	1.70	0.157	
Orientat*Margin	1	95.06	95.06	0.98	0.338	
Error	15	1455.94	97.06			
Total	63	30356.94				

9) Frustration (Mental Workload)

Source	DF	SS	MS	F	P	
Orientat	1	118.27	118.27	0.89	0.361	
Margin	1	28.89	28.89	0.17	0.690	
Subject	15	29513.98	1967.60	8.77	0.000	x
Orientat*Subject	15	1993.98	132.93	1.60	0.185	
Margin*Subject	15	2613.36	174.22	2.10	0.081	
Orientat*Margin	1	192.52	192.52	2.32	0.148	
Error	15	1242.73	82.85			
Total	63	35703.73				

10) Total Mental Workload (with the same weight)

Source	DF	SS	MS	F	P	
Orientat	1	45.14	45.14	0.46	0.506	
Margin	1	10.77	10.77	0.09	0.765	
Subject	15	24685.91	1645.73	10.97	0.000	x
Orientat*Subject	15	1456.81	97.12	1.53	0.208	
Margin*Subject	15	1742.75	116.18	1.84	0.125	
Orientat*Margin	1	193.38	193.38	3.06	0.101	
Error	15	949.19	63.28			
Total	63	29083.96				

11) Satisfaction

Source	DF	SS	MS	F	P	
Orientat	1	11.39	11.39	0.06	0.815	
Margin	1	28.89	28.89	0.23	0.635	
Subject	15	12369.61	824.64	2.96	0.012	x
Orientat*Subject	15	3000.86	200.06	4.46	0.003	
Margin*Subject	15	1848.36	123.22	2.75	0.029	
Orientat*Margin	1	58.14	58.14	1.30	0.273	
Error	15	672.11	44.81			
Total	63	17989.36				

CHI-SQUARE TEST RESULTS

1) Orientation Preference

The FREQ Procedure (Orientation Preference)

Orientation	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	8	50.00	8	50.00
1	8	50.00	16	100.00

Chi-Square Test for Equal Proportions

Chi-Square	0.0000
DF	1
Pr > ChiSq	1.0000

Sample Size = 16W

2) Margin Preference

The FREQ Procedure

Margin	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	10	62.50	10	62.50
1	6	37.50	16	100.00

Chi-Square Test for Equal Proportions

Chi-Square	1.0000
DF	1
Pr > ChiSq	0.3173

Sample Size = 16

3) Treatment Preference

The FREQ Procedure

Treat	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	3	18.75	3	18.75
2	4	25.00	7	43.75
3	6	37.50	13	81.25
4	3	18.75	16	100.00

Chi-Square Test
for Equal Proportions

Chi-Square	1.5000
DF	3
Pr > ChiSq	0.6823

WARNING: The table cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 16

4) Post-Experiment (Frustration)

The FREQ Procedure

Frust	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	2	12.50	2	12.50
1	14	87.50	16	100.00

Chi-Square Test
for Equal Proportions

Chi-Square	9.0000
DF	1
Pr > ChiSq	0.0027

Sample Size = 16

5) Post-Experiment (Usefulness)

The FREQ Procedure

useful	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1	6.25	1	6.25
1	15	93.75	16	100.00

Chi-Square Test
for Equal Proportions

Chi-Square	12.2500
DF	1
Pr > ChiSq	0.0005

Sample Size = 16

6) Post-Experiment (Perception)

The FREQ Procedure

perception	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	2	12.50	2	12.50
1	14	87.50	16	100.00

Chi-Square Test
for Equal Proportions

Chi-Square	9.0000
DF	1
Pr > ChiSq	0.0027

Sample Size = 16

SUPPLEMENTAL ANOVA RESULTS

1) ANOVA to test the effect of passages

Factor	Type	Levels	Values			
Passage	fixed	4	1	2	3	4

Analysis of Variance for Reading speed

Source	DF	SS	MS	F	P
Passage	3	5204	1735	1.39	0.255
Error	60	74976	1250		
Total	63	80180			

Analysis of Variance for Reading Comprehension

Source	DF	SS	MS	F	P
Passage	3	0.22500	0.07500	1.44	0.240
Error	60	3.12500	0.05208		
Total	63	3.35000			

Analysis of Variance for Proficiency

Source	DF	SS	MS	F	P
Passage	3	13703	4568	2.74	0.051
Error	60	99955	1666		
Total	63	113657			

Analysis of Variance for Perception of Ease of Reading

Source	DF	SS	MS	F	P
Passage	3	110.5	36.8	0.07	0.975
Error	60	30592.2	509.9		
Total	63	30702.7			

Analysis of Variance for Mental Demand

Source	DF	SS	MS	F	P
Passage	3	200.3	66.8	0.13	0.942
Error	60	30730.9	512.2		
Total	63	30931.2			

Analysis of Variance for Performance

Source	DF	SS	MS	F	P
Passage	3	656.2	218.7	0.34	0.795
Error	60	38410.9	640.2		
Total	63	39067.1			

Analysis of Variance for Effort

Source	DF	SS	MS	F	P
Passage	3	545.1	181.7	0.37	0.778
Error	60	29811.9	496.9		
Total	63	30356.9			

Analysis of Variance for Frustration

Source	DF	SS	MS	F	P
Passage	3	408.4	136.1	0.23	0.874
Error	60	35295.3	588.3		
Total	63	35703.7			

Analysis of Variance for Total Mental Workload

Source	DF	SS	MS	F	P
Passage	3	388.5	129.5	0.27	0.846
Error	60	28695.4	478.3		
Total	63	29084.0			

Analysis of Variance for Satisfaction

Source	DF	SS	MS	F	P
Passage	3	209.0	69.7	0.24	0.872
Error	60	17780.3	296.3		
Total	63	17989.4			

VITA

Young Seok Lee

EDUCATION

MS
Seoul National University of Technology, 02/2001, Dept. of Industrial Engineering
BS
Seoul National University of Technology, 02/2001, Dept. of Industrial Engineering

WORK EXPERIENCE

09/02 – 08/03 **Research Assistant**
"Usability Inspection, Testing, and Evaluation of Cell Phone User Interfaces" funded by Toshiba Corporation of Japan, Administered a number of focus groups to capture user requirements for the cell phone design. Designed and performed the entire usability testing procedure with a number of participants including people with disabilities. Analyzed quantitative data using SAS, qualitative data using video analysis to create design recommendation on the basis of universal design concept.

Grado Dept. of Industrial and Systems Engineering
Virginia Polytechnic Institute and State University, Blacksburg, VA

01/02 – 05/02 **Graduate Teaching Assistant**
ISE 3014 Work Measurement and Method Engineering
Taught laboratory sections, graded reports and homework assignments

Grado Dept. of Industrial and Systems Engineering
Virginia Polytechnic Institute and State University, Blacksburg, VA

08/01 – 12/01 **Graduate Teaching Assistant**
ISE 3614 Human Factors Engineering
Taught laboratory sections, graded reports and homework assignments

Grado Dept. of Industrial and Systems Engineering
Virginia Polytechnic Institute and State University, Blacksburg, VA

09/99 – 05/01 **Research Assistant**
"Visualization of Forging Processing Using Virtual Reality"
Funded by Korean Ministry of Science & Technology. Created a prototype of the Interactive Virtual Press using VRML. Also created user interface using Java EAI (External Authoring Interface).

Dept. of Industrial Engineering
Seoul National University of Technology, Korea

03/00 – 06/00 **Graduate Teaching Assistant**
Computer Aided Design -Pro/ENGINEER
Taught the course and laboratory sections with Pro/ENGINEER 2000i. Created and graded exams. Also graded reports and homework assignments.

Dept. of Industrial Engineering
Seoul National University of Technology, Korea

09/99 – 12/99 **Graduate Teaching Assistant**
Manufacturing System Engineering
Taught laboratory sections -"Virtual Manufacturing"- with Superscape-VRT,

Dept. of Industrial Engineering
Seoul National University of Technology, Korea

PUBLICATION

Lee, Y.S., Hwang, H. J., Jang, D.Y., & Kim J.P. (2001), Web-Based Visualization of Forging Operation by Using Virtual Reality Technique, *Industrial Engineering Interfaces*, Vol.14, No.1, 1-8.

CERTIFICATE and TRAINING

Certificate in Basics of Supply Chain Management of CPIM (Certified in Production and Inventory Management) from APICS (American Production and Inventory Control Society), 03/99

Manufacturing Training Course for CNC Lathe, Milling, Virtual Reality Robot and CIM, Denford Professional Training & Development Centre, Brighouse, West Yorkshire, England, 08/99

AWARDS and HONORS

08/01 – 08/03	Information and Telecommunication National Scholarship Program, The Ministry of Information and Communication, Republic of Korea
09/00	Seoul National University of Technology Scholarship (academic)
03/00	Seoul National University of Technology Scholarship (academic)
03/97	Seoul National University of Technology Scholarship (academic)

SKILLS

Languages: VRML, Java, HTML, JavaScript, ASP, C#, ASP.NET
Tools: Cosmo-World, JBuilder, JDK, Visual Studio.NET, Pro/ENGINEER. Flash

RESEARCH INTEREST

HCI, Cell phone User Interface Design, Digitizing Education, Electronic Books, Distance Learning, Usability, Industrial Application of Virtual Reality, Web3D