

A STUDY OF THE READABILITY OF ON-SCREEN TEXT

By

Eric Michael Weisenmiller

Dissertation submitted to the faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Curriculum and Instruction

Approved:

Dr. Mark E. Sanders, Chairman

Dr. James E. LaPorte

Dr. E. Allen Bame

Gail M. McMillan

Dr. Jimmie C. Fortune

July, 1999

Blacksburg, Virginia

A STUDY OF THE READABILITY OF ON-SCREEN TEXT

Eric Michael Weisenmiller

ABSTRACT

This study examined the readability of fonts. More specifically, it investigated how four different fonts effected both reading rate and reading comprehension. The typefaces Georgia, Verdana, (which, according to their designers, optimize on-screen readability) Times, and Arial (both designed for digital output to hard copy) were displayed as treatments both on a computer screen and on paper. The purpose of the study was to determine whether sans serif and serif typefaces optimized for on-screen viewing significantly improve reading rate and reading comprehension. Comparisons were made among the typefaces using a categorical independent variable postmeasure-only research design to determine the level of dependent variables (rate, comprehension). The group means of each of twelve treatment groups (N=264) were analyzed using analyses of variance to determine if either of the variables (presentation mode or font) had a statistically significant effect upon reading rate and/or reading comprehension of a sample taken from a population of subjects attending a midwestern state university. No significant difference was found among reading speed or reading comprehension scores of subjects tested who read text which was typeset in any of the four typefaces.

However, significant difference was found between the presentation modes used in the experiment.

Since it was found that 8-bit on-screen text was not significantly more readable than 600dpi text on paper, and 1-bit on-screen text was found to be significantly less readable than on-screen text and 600dpi text on paper, this research concludes that for purposes of ease of readability, on-screen text is better suited to be rendered as 8-bit on-screen text than 1-bit on-screen text. Also, the findings indicate that 8-bit on-screen text was not found to be significantly less readable than 600dpi text on paper. Also, due to the various typefaces currently being used in digital typography and the differing presentation media, further exploration of the readability of on-screen text should examine more fonts and screen display variables.

ACKNOWLEDGEMENTS

The work presented in this study is the culmination of countless hours of work and perseverance. The commitment, dedication, and moral support demonstrated by my committee, my professional colleagues, and my family have undoubtedly played a major role in seeing this project through to its completion. Thanks to all of you who have helped me endure one of the most challenging tasks of my lifetime.

I reserve special thanks for Dr. Mark Sanders, my committee chair. Thank you for giving me the opportunity to work with you as a graduate teaching assistant, for pushing me to complete my degree, and for the golf tips! I am most certain that the things you have taught me over the past four years will serve me well throughout my life.

I am also most appreciative of the efforts of my dissertation committee, Mrs. Gail McMillan, Dr. Jimmie Fortune, Dr. Allen Bame, and Dr. James LaPorte. Your advice, patience, and willingness to serve have helped me tremendously. Thanks also to the faculty, graduate students, and students in the Department of Industrial Technology at Illinois State University for your patience, guidance, and cooperation.

Finally, I say thank you to my family: Mike, Maggie, Gretchen, and Tim Weisenmiller. You've always supported me in your own special way. Thanks for the DNA, Mom and Dad!

TABLE OF CONTENTS

| | |
|---------------------------|------|
| ABSTRACT..... | ii |
| ACKNOWLEDGEMENTS..... | iv |
| TABLE OF CONTENTS..... | v |
| LIST OF TABLES..... | ix |
| LIST OF FIGURES..... | xiii |
| | |
| CHAPTER ONE..... | 1 |
| Introduction..... | 1 |
| Need for the Study..... | 5 |
| Purpose of the Study..... | 9 |
| Research Questions..... | 10 |
| Assumptions..... | 11 |
| Limitations..... | 11 |
| Definitions..... | 12 |
| | |
| CHAPTER TWO..... | 18 |
| Readability of Text..... | 18 |
| Typographic Research..... | 19 |
| Digital Typography..... | 27 |

| | |
|--|-------|
| Georgia and Verdana: Typefaces Optimized for the Computer Screen | 30 |
| | |
| Typography on the World Wide Web..... | 32 |
| How Humans Read..... | 37 |
| <i>Nelson-Denny Reading Test</i> | 40 |
| | |
| CHAPTER THREE..... | 42 |
| Research Objectives..... | 42 |
| Statistical Hypotheses..... | 46 |
| Research Design..... | 48 |
| Population and Sample..... | 52 |
| Test Instrument..... | 54 |
| Testing Procedures..... | 55 |
| Administration of the Test..... | 58 |
| Variables..... | 59 |
| Control of Extraneous Variables..... | 60 |
| Statistical Procedures and Data Analysis..... | 62 |
| | |
| CHAPTER FOUR..... | 63 |
| Introduction..... | 63 |
| Analysis of Data..... | 64 |
| Group Homogeneity..... | 66 |

| | |
|--|-----|
| Testing the Null Hypotheses..... | 67 |
| Reading Rates..... | 68 |
| Reading Comprehension..... | 71 |
| Findings..... | 75 |
| Other Analyses..... | 88 |
| Analysis Summary..... | 93 |
| | |
| CHAPTER FIVE..... | 94 |
| Background..... | 94 |
| Purpose..... | 96 |
| Procedure..... | 96 |
| Findings..... | 98 |
| Conclusions..... | 101 |
| Observations..... | 101 |
| Discussion..... | 103 |
| Recommendations..... | 107 |
| Implications for Practice..... | 110 |
| | |
| REFERENCES..... | 111 |
| | |
| APPENDIX A: Typography Primer..... | 123 |
| APPENDIX B: Instructions for the Reading Test..... | 128 |

| | |
|---|-----|
| APPENDIX C: Courses in the Department of Industrial Technology from which the subjects in this subject were sampled..... | 133 |
| APPENDIX D: Answer Sheet..... | 134 |
| APPENDIX E: Times Displayed as 1-bit and 8-bit Text..... | 136 |
| APPENDIX F: Arial Displayed as 1-bit and 8-bit Text..... | 137 |
| APPENDIX G: Georgia Displayed as 1-bit and 8-bit Text..... | 138 |
| APPENDIX H: Verdana Displayed as 1-bit and 8-bit Text..... | 139 |
| VITA..... | 140 |

LIST OF TABLES

| | |
|---|----|
| Table 1: Treatment Groups..... | 45 |
| Table 2: Factorial Design of the Study Illustrating the Categorical Independent Variable Post-Measure Only Design..... | 50 |
| Table 3: Factorial Design 3 X 4 Analysis of Variance (ANOVA)..... | 51 |
| Table 4: Group Means (Standard Deviations) for Reading Rate Test..... | 65 |
| Table 5: Group Means (Standard Deviations) for Reading Comprehension Test | 66 |
| Table 7: Summary of Two-way ANOVA of Reading Rate Scores..... | 69 |
| Table 8: One-Way ANOVA for Reading Rate Across Three Levels of Presentation Mode..... | 69 |
| Table 9: Descriptive Statistics for Reading Rate Scores across Three Levels of Presentation Mode..... | 70 |
| Table 10: Tukey's Pairwise Comparisons for Reading Rate Scores Across Three Levels of Presentation Mode..... | 70 |
| Table 11: Summary of Two-way ANOVA of Reading Comprehension Scores for Reading Comprehension..... | 72 |
| Table 12: One-Way ANOVA on Comprehension Scores across Three Levels of Presentation Mode..... | 73 |
| Table 13: Descriptive Statistics of Comprehension Scores across Three Levels of Presentation Mode..... | 73 |

| | |
|---|----|
| Table 14: Tukey's Pairwise Comparisons of Comprehension Scores across Three Levels of Presentation Mode..... | 74 |
| Table 15: One-Way ANOVA of Reading Rate Scores Across Four Levels of Typeface Variation as They Were Displayed as 1-bit Text on a Computer Screen..... | 76 |
| Table 16: Descriptive Statistics of Reading Rate Scores across Four Levels of Typeface Variation as They Were Displayed as 1-bit Text on a Computer Screen..... | 76 |
| Table 17: ANOVA of Reading Comprehension Scores across four levels of typeface variation as displayed as 1-bit text on a computer screen.... | 78 |
| Table 18: Descriptive Statistics of Reading Comprehension Scores across Four Levels of Typeface Variation as Displayed As 1-bit Text on A Computer Screen..... | 79 |
| Table 19: One-Way ANOVA of Reading Rate Scores across Four Levels of Typeface Variation as They Were Displayed as 8-bit Text on a Computer Screen..... | 80 |
| Table 20: Descriptive Statistics of Reading Rate Scores across Four Levels of Typeface Variation as They Were Displayed as 8-bit Text on a Computer Screen..... | 81 |
| Table 21: ANOVA Of Reading Comprehension Scores Across Four Levels Of Typeface Variation As They Were Displayed As 8-bit Text On A Computer Screen..... | 82 |

| | |
|---|----|
| Table 22: Descriptive Statistics of Reading Comprehensions Scores across Four Levels of Typeface Variation as They Were Displayed as 8-bit Text on a Computer Screen..... | 84 |
| Table 23: One-Way ANOVA of Reading Rate Scores across Four Levels of Typeface Variation as They Were Displayed as 600dpi Text on Paper | 85 |
| Table 24: Descriptive Statistics of Reading Rate Scores across Four Levels of Typeface Variation as They Were Displayed as 600dpi Text on Paper | 85 |
| Table 25: ANOVA of Reading Comprehension Scores across Four Levels of Typeface Variation as They Were Displayed as 600dpi Text on Paper | 87 |
| Table 26: Descriptive Statistics for Reading Comprehension Scores across Four Levels of Typeface Variation as They Were Displayed as 600dpi Text on Paper..... | 87 |
| Table 27: One-Way ANOVA Comparing Reading Rate Scores across Three Presentation Modes..... | 88 |
| Table 28: Descriptive Statistics for Reading Rate Scores across Three Presentation Modes..... | 89 |
| Table 29: Tukey's Pairwise Comparisons for Reading Rate Scores across Three Presentation Modes..... | 89 |

| | |
|---|----|
| Table 30: One-Way ANOVA Comparing Reading Rate Scores across the Three Presentation Modes in Times..... | 91 |
| Table 31: Descriptive Statistics for Reading Rate Scores across the Three Presentation Modes in Times..... | 91 |
| Table 32: Tukey's Pairwise Comparisons for Reading Rate Scores across the Three Presentation Modes in Times..... | 92 |
| Table 33: Treatment Groups..... | 98 |

LIST OF FIGURES

| | |
|--|----|
| Figure 1. The “look” of bilevel fonts and grayscale fonts..... | 12 |
| Figure 2. The Look of Grayscale Typefaces..... | 13 |
| Figure 3: Verdana, Arial, Georgia, and Times..... | 32 |
| Figure 4. The results of the reading rate test across the different levels of typeface with respect to the mode of presentation..... | 71 |
| Figure 5. The results of the reading comprehension test across the different modes of presentation with respect to the four levels of typeface variation..... | 75 |
| Figure 6. The results of the reading rate test across the four different levels of typeface as they were presented using 1-bit text on a computer screen..... | 77 |
| Figure 7. The results of the reading comprehension test across the four different levels of typeface as they were presented using 1-bit text on a computer screen..... | 79 |
| Figure 8. The results of the reading rate test across the four different levels of typeface as they were presented using 8-bit text on a computer screen..... | 82 |
| Figure 9. The results of the reading comprehension test across the four different levels of typeface as they were presented using 8-bit text on a computer screen..... | 83 |

Figure 10. The results of the reading rate test across the four different levels of typeface as they were presented on 600 dpi paper output.....86

Figure 11. The results of the reading comprehension test across the four different levels of typeface as they were presented on 600 dpi paper output.....88

Figure 12. The results of the reading comprehension test across the different levels of typeface with respect to the mode of presentation.....90

CHAPTER ONE

Introduction

The nature of typography is multi-dimensional. It is likened by many to an art and to a technology by others. Regardless, none can dispute technology's role in typography's evolution. Ever since the days of Johannes Gutenberg, the distribution of information—and, in turn, knowledge—has relied upon the technological innovation of the process of reproducing textual matter. Many point to Gutenberg's invention of movable type in the mid-1400s as one of the most significant inventions in history. However, the current age is also an era of remarkable advances in the technology of text-based communication—so much so that it has become known as the "information age".

Throughout history, it seems that issues of typography have weighed heavily upon the balance of aesthetics versus technology—or form versus function. At some points throughout history mankind has been a slave to the technology of text, at others it has assumed the role of master. Regardless, typography's purpose is to augment meaningful communication. Moreover, it serves to graphically represent the author's message while simultaneously portraying an aesthetic value or tone. Ultimately, however, typography conveys both overt and covert meaning. Will-Harris (1998), a recognized modern typographic expert and publisher of the online journal *Typofile*, expresses the essential function of typefaces by characterizing their uniqueness among other

means of information transmission by amplifying the fact that text “transmit [s] complex intellectual and emotional messages in a very concise and precise way” (1998).

Prior to our decade, the majority of text-based reading took place between reader and paper as one’s eyes scanned the lines of a printed page. Nowadays, with the rapid proliferation of personal computers, an increasing amount of text is being read directly from CRTs, LCDs, and TVs. Although many clinging to the ways of the past, the shift from the traditional “print paradigm” to on-screen reading mirrors an overriding societal momentum shift toward a digital culture which began with the proliferation of personal computers in the early 1980s with the advent of the affordable desktop computer.

A result of what many refer to as the “desktop publishing revolution” is the increased ease, accessibility, and affordability of publishing capability. DiNucci (1995) recognizes two divergent paths taken in the type design resulting from the development of widespread on-screen design. One faction of the type design community is currently attempting to remain true to the traditions of fine typography by exploiting the innovative digital tools of the present, while the other camp is seeking to exploit the possibilities of the characteristics of the new media.

Since its advent, desktop publishing in concert with networked communication advances (most noticeably, the World Wide Web, or WWW) has led to a boom in electronic publishing. It has also led to a rapid shift in how the

“printed” word is presented and, in turn, viewed by the reader. More and more, textual information (personal correspondence, educational/ informational material, books, magazines, newspapers, etc.) is being designed for/obtained for computer screens instead of from traditional print media. Such a rapid technological shift naturally calls us to periodically question our innovations and their effects on established rules of the past.

In 1991, a study of the effects of desktop publishing on the craft of typography, Cartwright found that much to traditional typographers' dismay, they "perceive a general decline in the knowledge of their craft." Powerful and inexpensive desktop computers with elaborate typographic tools, which have become widely accessible to novices, were cited as the cause for this decline. In that same year, the creation of the World Wide Web (WWW) by Tim Berners-Lee and his development team at the European Laboratory for Particle Physics (CERN) marked a monumental point in the history of communication. Their development of hypertext markup language (HTML), the enabling code behind WWW pages, allowed "browsers" to see richly formatted documents with nice-looking fonts, emphasis, and text alignment via the Internet on pages displayed on their computer screens (Pfaffenberger, 1996, p. 33). This innovation both sustained and extended the boundaries of electronic publishing and created a boom in the medium during the mid-1990s. Suddenly, global publishing capabilities were in the hands of practically anyone who had access to a personal computer. Both events, in effect, have placed global publishing

capabilities that were once the singular domain of printers and publishers into the hands of novices.

Applications of electronic publishing, many facilitated by the Internet, have now come to affect many facets of everyday life. Education is undoubtedly one of the major beneficiaries of the growth of networked communication as libraries of printed material have been digitized and made easily accessible via electronic communication devices. Distance learning opportunities have increased many-fold during the latter part of the 1990s as a result of widespread access to the global information conduit of the Internet. In addition, electronic commerce, personal communication, and political propaganda now have a new venue. When one considers the enormity of the growth of use of recent technological innovation in the area of communication technology, the need for a reexamination of the basic foundations of this new electronic media become strikingly evident.

Previous readability research has focused upon the effects of typographic and page-layout variables on reading rate and comprehension, as well as mental and physical human factors. As Holmes (1986) points out, prior research in the area of functionality of textual information—or readability—concluded that a combination of reader and typographic variables affect reading rate and reading comprehension. Less conclusive, however, are reasearchers findings related to the display of information on new computer display media such as CRTs and LCDs. Although research exists which

examines the differences between the readability of textual material on-paper and on-screen, the rapid evolution of digital typographic tools and display technologies call researchers to periodic re-examination in order to guard against inefficiency.

A review of the literature surrounding the study of readability indicates a scarcity of research which substantiates the advantages of using screen-optimized typefaces for the purposes of textual display on a computer screen. These commercially available typefaces are touted as "more readable" for the purposes of screen display of text. Most commonly they are employed in the creation of web pages, electronic books (e-books), or e-mail messages. However, some research has indicated their suitability for use in the on-screen display of Adobe Portable Document Files (PDFs) (Mather, 1997).

Need for the Study

As humanity progress into the "information age", a paradigm shift in the area of information delivery has become apparent. The dominant medium of textual information delivery is currently in the midst of a transition from print to computer display. Evidence of this shift abounds. The Internet Society (1998) reported that the mere 100,000 sites which existed on the WWW in January 1996 had grown to 2,215,195 sites by April 1998 (an increase of 2,200 percent). Merely citing statistics is not to boldly suggest that soon *all* textual information will be delivered digitally onto a computer display—the printed word

still flourishes. However, when one considers that over 41.5 million adults in the United States (PR Newswire 1998) use the Internet daily—and one takes into consideration that text makes up the majority of information on most pages—how people interact with computer displays deserves a measure of scrutiny.

In light of this evidence, it becomes the responsibility of researchers to continue to examine both the ergonomic issues and behavioral issues related to the display of text as they have done in the past (Holmes, 1986). However, as the media for textual display transitions from print to computer display, researchers must take into consideration the dynamic dimensions offered by new media. As High (1997) pointed out, researchers ought to examine the factors unique to new media and adjust our typographic conventions to fit a new model, (while avoiding) the temptation to superimpose...old typographic guidelines, which were refined over the years for print media on the digitally produced typographic layouts of the present and future.” Reinforcing these time-honored guidelines by default would be limiting. High (1997) elaborated upon this danger by stating:

“...just as we have the QWERTY keyboard 'standard' today based upon a deliberate engineering choice to slow typing rate and thus prevent mechanical failure of typewriters, we are making questionable decisions today which will set patterns of development and 'standards' for future

generations. In other words, do we have our paradigms on straight? Especially when it comes to reading in the digital dimension, the potential for harm is quite real and ultimately predictable.”

Typefaces designed for the screen are a relatively new innovation. They represent an attempt to compensate for the lack of resolution of the bitmaps of computer displays. They also attempt to compensate for the fact that the majority of fonts available today were created to be digitally typeset but to be read only after being printed on paper. These older fonts, in effect, forced one to read “type designed in another era for another medium” (Will-Harris 1998). Changes in the generally accepted rules and conventions of a profession so steeped in tradition will surely be difficult, however these changes are a necessity if progress in typography is to be made. This has increasingly become a concern in relation to computer software as more people have increasingly begun to read from computer displays. On-screen fonts, on the other hand, are fonts that have been designed from the start to optimally render typographic characteristics (characteristics such as x-height, letter spacing, and serifs that increase legibility) on the bitmapped screens of computers. Thus attempting to overcome an obvious technical shortcoming of on-screen text displays.

As one of the initial companies recognizing a need for the development of fonts designed specifically for on-screen viewing, Microsoft Corporation

commissioned the development of several screen fonts in the Summer of 1994 to be used in their *Windows 95* operating system (Will-Harris, 1998). The overwhelming popularity of the *Windows 95* operating system, coupled with the company's initiative to distribute these fonts for free threaten to place these fonts as the de facto standard on the WWW.

Considering Microsoft's dominance and influence which have placed their products as "standards" in today's personal computing society, it would be natural to question the validity of their innovations (which have historically become tantamount to technological mandates). Other companies including Adobe Systems, Bitstream, and Monotype have entered the market for on-screen fonts. However, the widespread proliferation of Microsoft's free fonts is inevitable due to the overwhelming use of Microsoft products in the personal computer arena.

In 1998, Microsoft Corporation Chairman Bill Gates said his company was working on new tools aimed at improving web-page readability. Speaking with reporters after his Comdex keynote, Gates said, "Web layout is out of Microsoft's control, but the company is working on rich text controls in Windows that will improve Web readability." Gates bemoaned layouts that make reading on the Web difficult, such as wide columns and scrolling pages. Microsoft developers plan to focus on these and other readability issues, Gates said (Will-Harris, 1998).

Previous readability research has examined printed typography, and typography designed for print as it has been rendered on screen. In the majority of studies it has been shown that reading performance of subjects who read text presented on paper is much better than that of subjects reading from alternative presentation media. However, there exists little research which examines the legibility of typography that was expressly designed for the screen. There also exists virtually no research which examines the readability of text rendered using Adobe's Portable Document Format.

Purpose of the Study

This study examined the construct of readability, or how typographic characteristics affect reading performance as measured by reading rate and reading comprehension. The purpose of this study was to determine empirically whether the typefaces Georgia and Verdana (which are optimized for on-screen viewing) significantly improve reading performance as measured by reading rate and/or reading comprehension as compared to the designed-for-print typefaces Times and Arial—both of which are digital fonts designed specifically for text-on-paper. In addition, all typefaces were tested and compared as they appear on a computer display—both as 1-bit text and as 8-bit text—and on paper as 600dpi laser-printed output.

Research Questions

1. Was there significant difference in reading rate scores of subjects who read text passages set in four different typefaces (Times, Georgia, Arial, and Verdana) displayed as 1-bit text on a computer screen?
2. Was there significant difference in reading comprehension scores of subjects who read text passages set in four different typefaces (Times, Georgia, Arial, and Verdana) displayed as 1-bit text on a computer screen?
3. Was there significant difference in reading rate scores of subjects who read text passages set in with four different typefaces (Times, Georgia, Arial, and Verdana) rendered at 600dpi on paper?
4. Was there significant difference in reading comprehension scores of subjects who read text passages set in with four different typefaces (Times, Georgia, Arial, and Verdana) rendered at 600dpi on paper?
5. Was there significant difference in reading rate scores of subjects who read text passages set in four different typefaces (Times, Georgia, Arial, and Verdana) displayed as 8-bit text on a computer display.
6. Was there significant difference in reading comprehension scores of subjects who read text passages set in with four different typefaces (Times, Georgia, Arial, and Verdana) displayed as 8-bit text on a computer display?

Assumptions

The following assumptions were made with respect to this study:

1. Subjects chosen for this study were representative of the Industrial Technology student population at Illinois State University.
2. Since the *Nelson-Denny Reading Test* contains textual passages considered appropriate for students between the 9th to the 16th grade level, all subjects' reading ability was assumed to be within the range of the 9th to the 16th grade level.

Limitations

This study was conducted under the following limitations:

1. Given that the sample for this study was taken from the student population at Illinois State University, the results can only be truly be generalized to the Illinois State University Department of Industrial Technology population. However, a hypothetical generalization could be applied to university populations with similar demographics.
2. The accuracy of the reading rate and reading comprehension scores was limited to the accuracy of the standardized test instrument (*The Nelson-Denny Reading Test* - Form G) as it is adapted for this research. This test limits the subject to a twenty-minute time limit.

Definitions

Typographic terms specific to this study are operationally defined below:

1-bit text. Bilevel text in one bit color (i.e. either black or white). That is, the letterforms are aliased, or stair-stepped (jaggy) at their curved edges as if conforming to a bitmap grid [See *Figure 1*].



Figure 1. The “look” of 1-bit (bilevel) fonts (400% Magnification)

8-bit text. 8-bit text is text displayed using 8-bit grayscale color. That is, the letterforms are anti-aliased, or dithered at their curved edges as if to smooth the edges of the natural "jaggy" appearance of text rendered on a bitmap. Also known as grayscale typefaces, they are typefaces that are “generated by filtering and re-sampling high-resolution bilevel master characters” (O’Regan, K., Bismuth, N., Hersh, R.D., & Pappas, A.,1998). [See *Figure 2.*]

Arial Eight-Bit Text

Figure 2. The Look of Grayscale Typefaces (400% Magnification)

Bitmap fonts. Bitmap fonts are fonts designed for low-resolution-screen display (Cavanaugh 1995). Bitmapped fonts fall into two basic categories: bilevel and grayscale. According to O'Reagan et al. (1998), bilevel fonts are rendered in two levels of intensity (usually black and white), whereas grayscale fonts are "generated by filtering and re-sampling high-resolution bilevel master characters". Fonts on the web are rendered as bilevel fonts.

Contrast. Contrast is "a dynamic polarity that helps to clarify a graphic idea. Contrast is a force of visual intensity and as such it simplifies the process of communication" (Berryman, 1984, p. 28).

Counter. Counters are the non-printing areas around characters and inside the loops of text characters (Prust, 1989, p. 69).

CRT. cathode ray tube

Font. A font is a complete assortment of upper and lower case characters, numerals, punctuation, and other symbols of one typeface (Beach, 1992, p.104)

Georgia. Georgia is a typeface designed specifically for on-screen viewing by Mathew Carter, a well respected type designer with several typeface designs to his credit. This typeface along with Verdana was commissioned by Microsoft to optimize two popular print-optimized typefaces Times and Arial for the CRT

Hints (or "hinting"). "Modern scalable digital typefaces incorporate hints to tell the rasteriser how best to render the font at given point sizes. Good hinting is crucial to accurate rendering (or grid fitting) on bitmapped displays for small point sizes. Several visual artifacts can be introduced in displaying unhinted fonts on bitmapped displays, including uneven colour, inconsistent spacing, weight, and alignment, and poor symmetry. Hinting can greatly alleviate these problems, and it is fair to say, the quality of a computer font can be measured in the quality of its hints" (Mather, 1997, p. 10).

Legibility. Legibility "is concerned with type design and the visual shape of individual type characters. Legibility is the rate with which a type character can be identified" (Berryman, 1984, p. 28).

Readability. Readability is "the ease of reading a printed page or message. It refers to the arrangement of type(s). Readability involves design of the total visual entity, the complex interrelations among type, symbols, photos, and illustrations" (Berryman, 1984, p. 28).

Resolution. Resolution is the sharpness of an image on film, paper, CRT, disk, tape, or other medium (Beach, 1993, p.204).

Portable Document Format (PDF). " Acrobat Portable Document Format (PDF) is the open de facto standard for electronic document distribution worldwide. PDF is a universal file format that preserves all of the fonts, formatting, colors, and graphics of any source document, regardless of the application and platform used to create it. PDF files are compact and can be shared, viewed, navigated, and printed exactly as intended by anyone with a free Adobe Acrobat Reader. "

Print. Print is material containing text, illustrations, and/or halftone photographs reproduced in some quantity by means of implementation of a process through which ink is impressed upon a substrate (especially paper).

Serif font. A serif font contains short strokes or thickened tips at the ends of the character strokes of its letters (Prust, 1989, pp.70).

Sans serif font. In a sans serif font the characters are without serifs (Prust, 1989, p. 75).

Text. Text refers simply to "written discourse (aggregates of words) in printed form...whether that text is reproduced on paper or in electronic signals on a cathode ray tube" (Jonassen, 1982, ix).

Typeface. A typeface is a set of characters with similar design features and weight.

Typography. Typography is "the application of design principles to the setting of type," or, more simply: "choosing and using type" (Berryman, 1984, p. 28).

Verdana. Verdana is a typeface designed specifically for on-screen viewing by Mathew Carter, a well-respected type designer with several typeface designs to his credit. This typeface along with Georgia was commissioned by Microsoft to optimize two popular print-optimized typefaces Times and Arial for the CRT

X-height. X-height is the measure in points of the distance from the top to the bottom of lowercase letters—excluding ascenders and descenders (Prust, 1989, p. 69).

CHAPTER TWO

Readability of Text

The concept of readability optimization is not new; several of its facets have been examined and tested over the past fifty years. Research has shown that readability can vary in accordance with certain specific typographic variables. It has been shown, for example, that characteristics such as serifs (or the lack thereof) can either speed or slow reading rates of text in print (Taylor, 1990). However, the overwhelming majority of this research has focused upon the readability of text in print. This stems from the fact that most textual content prior to the past decade has existed in the form of print. These studies have examined such factors as the effects of typeface (e.g. serif versus sans serif typefaces), letterspacing, line spacing (or leading), justification contrast, resolution, inverted text, mechanically-tinted backgrounds, size, type style, letterspacing, and page layout. These typographic variables have been tested in order to determine various effects upon the reader. Chief among these variables are reading rate and reading comprehension.

The readability of serif and sans serif fonts has been a lingering point of contention among typographers and graphic designers. Although many studies have compared the legibility of serif font and sans serif fonts in print (Taylor, 1990; Kraveuttske, 1994), and some research exists on the reading rate of text displayed on various types of CRTs, there is currently little scientific research

which examines the legibility of screen optimized serif and sans serif fonts designed specifically for representation as hypertext on CRTs. With their introduction in 1995, these fonts represent a relatively recent innovation in typographic technology which should bear scientific examination.

Typographic Research

Although digital audio and digital video are pushing the limits of available bandwidth, and promise to play a major role in the future of networked communication, the most common component of the Internet is its most basic—text. The transmission of text-based material constitutes the majority of the material on the Internet and is, like print material, the prevalent mode of information transfer from page to reader. Throughout the ever-expanding evolution of networked modes of communication, “text” has remained a constant thus far.

Text refers simply to “written discourse (aggregates of words) in printed form...whether that text is reproduced on paper or in electronic signals on a cathode ray tube” (Johnassen, 1982, ix). Nowadays text as it is used in many forms of electronic publishing is no longer one-dimensional in nature as it had existed for centuries. Therefore, with technological advances in typographic tools, text can be more effectively designed, regardless of the medium through which it is transmitted” (1982, x).

Jonassen (1982) asserted his belief that despite the paradigm shift from print to on-screen type, text would prevail as a major form of recorded communication for the foreseeable future. Therefore, it is worthy of pursuing ways that text can be more effectively designed, regardless of the medium through which it is transmitted” (1982, x). However, technical restraints have limited designers by placing many design specifications in the hands of those who view (or browse) documents. This has, in effect, tied the hands of many designers whose formatting specifications can be easily over-ridden by the viewer (Bond, 1998). However, in the past three years, designers of WWW-based material have progressively wrested control over the design of their documents. Technological advances in the capabilities of HTML code have increased layout and typographic parameters of WWW pages. Embedded fonts, cascading style sheets, OpenType, and the tag are all examples of innovations that have attempted to ensure that the author/designer(s) retain(s) command of the design of their pages—as opposed to browser.

The emergence of a digital age has altered the manner in which many of us read. Instead of reading the printed word exclusively, many computer users do a major portion of reading from computer displays. Just as the textual design considerations of the past applied to paper media, examinations of the textual design of the future demand examination of digital display media. Reynolds (1996) indicates a need for this type of research by stating “...now that

desktop publishing systems have brought electronic typesetting within the reach of those with little or no knowledge of basic design principles [,] legibility as an issue is perhaps more important now than ever before.” This study focused upon text displayed on a cathode ray tube (CRT). It sought to examine physical attributes or, as Jonassen puts it “the technology of text [which is] the application of a scientific approach to text design”.

Research studies which compare paper and computer screen readability most often show that text displayed on computer (or video) screens is less readable than paper (Gould, 1981; Gould & Grischkowsky, 1984; Haas & Hayes, 1985; Kruk & Muter, 1984; Wright & Lickorish, 1983). These studies suggest that there are many factors that could potentially lead to improvements in screen readability.

The typographic design factors associated with presenting text both on-screen and on traditional paper media are numerous and often interrelated. Among the most common typographic factors that are examined in an experimental research context include: typeface (e.g. serif versus sans serif typefaces), letterspacing, line spacing (or leading), justification contrast, resolution, inverted text, mechanically-tinted backgrounds, type size, type style, and letter spacing. Other research has examined the media or display conditions upon which text is displayed as well as demographic factors of the readers themselves. Due to the sheer number and constantly changing nature of variables at work in the interaction between human and text, the research in

this area often serves as a starting point from which to build. Seemingly, each new innovation in typographic control and media becomes a new variable to consider in the visual information transfer from text to reader.

Canary (1983) experimented with typeface variations including (a) set width (standard, condensed, and extended), (b) print weight (regular, medium, and bold), and (c) leading (one-point and two-point). Although the study cited significance of the experiment in its conclusion that “specific features of typography can influence effectiveness of the reading process, statistical significance was only found in extended letter width (or set width) treatments being determined less legible ($p < .05$) than standard width or condensed widths across the two point weights and variations in leading. These findings are consistent with what digital typographers believe to be true about legibility optimization of fonts being displayed on CRTs. For example, the design of Verdana incorporates extended characters with extra letterspacing. This is essential to the legibility of the font as it appears on-screen because as Matthew Carter, the designer of Verdana explains, “pre-existing printer fonts are spaced for paper, not the screen, so they suffer on-screen. In Verdana, it’s the regularity of the spacing that’s just as important as the positive parts of the letterform” (Carter cited in Will-Harris, 1998). This illustrates the differences in design features of type displayed on either medium, screen and print. Similarly, Turner (1982) examined the degrees of legibility in respect to the amount of minus-letterspacing of the body type of certain textual passages that were

typeset in either a serif or a sans serif typeface. Minus-letterspacing is defined in this study as body type that is reduced below "normal" levels. Normal levels in this case being based upon a typeface with 18-unit em. Significance surfaced in favor of minus spacing as opposed to normal spacing. This finding refers to text in print. It also runs contrary to what the designers of on-screen fonts theorize is optimal for text displayed on a bitmap. Bitmaps being restricted to fixed resolutions.

Holmes (1986) studied the effects of both resolution and formatting dot matrix print with respect to formatting of text (fully justified and left justified). Dot matrix print represents low-resolution text-on-paper popular in the late 1970s and early 1980s. Holmes found that neither resolution nor formatting made any significant difference in the reading rate or reading comprehension scores of the subjects tested when compared with different quality levels of dot-matrix print output on-paper. However, Journa (1995) found that display resolution (image quality) is, in fact, a function of reading rate. In other words, as image quality increases, reading rate increases. This was true regardless of medium - CRT or print. The reason for the seemingly contrary results is probably due to the increments in quality differential in the two studies. That is, in Holmes study text in print was compared at two quality levels which were only slightly different, but Journa found significance in image quality in both print and on-screen text which had a wider range of image quality from best to worst. Journa also found that as quality increased, so did reader preference.

Taylor (1990) points out in his study, “The Effect of Typeface on Reading Rates and the Typeface Preferences of Individual Readers,” that on the basis of personal preference, readers chose the sans serif typeface more often than they chose the serif typeface. He concluded “designers and developers of texts and technical materials can use these conclusions as a means of preparing more efficient and effective documents. They can move the reader along with a serif typeface and then slow them down with a sans serif area that they are drawn to by a preference for that typeface.” Although reader preference is an issue of importance in designing fonts for the WWW, it can—as subjective measures wont to do—detract from the issue of real importance which is legibility. These issues are strictly objective in nature.

Tullis, Boynton, and Hersh’s (1995) study of the readability of fonts in the Windows environment found that among twelve different size/font variations tested, ranging in size from 6.0 point to 9.75 point, the font Arial 7.5 and any Small Fonts be avoided due to their lack of readability. Subjective data led to the suggested use of “Arial 9.75” or MS Sans Serif 9.75. As compared with all fonts in point sizes ranging from 6.0 to 9.75, optimal reading rate and accuracy was found in 8.25, 9.0, and 9.75 point fonts - except for MS Sans Serif 8.25. These findings argue for fonts that are large enough to be read on limited resolution CRTs.

Le Rohellec, J., Brettel, H., and Glassar, J. (1996) used a visual performance measure (as opposed to a performance test) to measure the

legibility of characters displayed on CRTs under different lighting conditions. Their aim was to determine the role of varying degrees of illumination in display legibility. The results of the Gerig, Nibbelink, and Hoover (1992) study did not indicate statistical significance of the influence of type size on reading comprehension as measured by standardized test forms. Their study addressed the origins of the textbook typographic conventions by questioning first how type-size conventions come into being, and second, is student achievement enhanced because of current typographic conventions. In the course of their research they challenged the British Association for the Advancement of Science's assertion that the size of the typeface is the most important factor in the influence of the display of text in books upon vision. However, as the authors point out, the size of the typeface is often equated with age. This convention, begun early in the century, often mandated larger type for younger children and, in turn, progressively smaller type for older, more advanced students. However, citing Gillard (1923) the authors point out that "children are not so greatly affected as adults by changes in size of type." This fact, they explained, was exposed by early twentieth century research findings (and the reluctance by many over-forty teachers/adults to change to bifocals).

Uhl (1937) points out that "the eye reaches maturity earlier than any other organism in the body" (p.28). This led Gerig, Nibbelink, and Hoover (1992) to almost entirely discount the probability of the negative influence of small type on the reading and/or visual development of young students. It has been

generally held since the technological advances of the 1920s and 1930s—advances that allowed researchers to more precisely examine the factors of eye movement—that changes in size of type have only a miniscule affect over reading rate. Buckingham (1931) points out the number of words recognized at each separate visual fixation only slightly increases despite the doubling of the point size.

Gerig, Nibbelink, and Hoover (1992), Gillard (1923), Buckingham (1931), and Uhl (1937) concur on the fact that increasing any type size above 12-point does not benefit the reading rate of school-aged children or those with corrective vision. They focused their research upon the effects of typographic formatting variables common to standardized tests on "reading comprehension".

Studies by the National Council for the Study of Education have shown that shorter line lengths resulting from setting type in two columns is definitively superior to setting type in one column (Buckingham 1931, p. 121). According to their 1992 study which examined the relationship between type size, line length, and line spacing, and the readability of text, 12-point type was read slightly more efficiently than 14-point type. Also, 18-point type was tested, but it was read much slower than either 12-point or 14-point type. The study found no advantage to using larger type sizes. Uhl's (1937) findings backed-up what Buckingham had concluded. He found that eye fatigue among young children reading passages of text was independent of type size.

Digital Typography

As an increasing amount of type will be set for on-screen reading in the future, digital typographic specifications should be considered. However, as has been stated earlier, the nature of digital typography is quite different than its print predecessor. CRTs often do not offer the degree of resolution available with most printing processes. In order to design on-screen typography, one must think in pixels. Pixels are the small, block units of a screen's bitmap which compose the strokes of on-screen type. When text on a CRT is tested, it takes on a number of different variables than text in print. Text rendered on-screen has a much lower resolution than printed type. Whereas text in print can appear in high resolution, on-screen text appears as low-resolution bitmaps. Screen resolutions of computers running the Windows 95 operating system (100 ppi) and the Macintosh operating system (72 ppi) are exceeded by nearly 25 percent by an a relatively poor-quality print from a fax machine set in normal mode (Will-Harris, 1988).

Will-Harris notes that size of typeface is a key variable in readability. He suggests that 14-16 point type is significantly easier to read than smaller type, while cautioning against any point size set at 10 point or smaller. Research agrees that some fonts cannot be read as effectively, especially fonts smaller than 7.5 points on a 100 ppi display (Tullis, et. al., 1995). Size weight is also considered a key variable in typeface selection. Will-Harris (1998) observes

that chunky, heavier typefaces often tend to “block up”—that is, the empty space that usually differentiates the strokes of characters is often reduced or eliminated. This is due to the limited resolution of CRTs.

Will-Harris cites Chuck Bigelow’s suggestions for choosing an on-screen typeface (Chuck Bigelow is the co-designer of the Lucida family of typefaces with Kris Holmes): (a) consider a sans serif typeface “When printed, the serifs on typefaces are only a tiny percentage of the typeface’s design. But on-screen, in order to display the serifs using the limited number of available pixels, they take up a much bigger proportion of available information than they do on a printed page. Serifs should be small things - but on-screen they become big - no longer visual cues but noise - distracting chunks of interference,” (b) “hinting of individual fonts can be as important as the typeface design itself,” (c) choosing a typeface with a large x-height. Due to the fact that in lower-case letters the majority of the readability recognition comes from visual information conveyed from the forms within the x-height, typefaces with larger x-heights naturally tend to be more legible. Bigelow suggests choosing an on-screen typeface with an “x-height that is one pixel larger than half the body size - so a 12-point typeface would have an x-height of 7 pixels,” and (d) adding extra tracking between characters. Bigelow elaborates: “Collisions between characters becomes very annoying on-screen - when two characters touch even by one pixel you get a lot of noise in the tangle of shapes”. Mather (1997) concurs with this advice with his conclusion that the proportion of x-

height to character width of lowercase characters as well as adding width can improve their appearance on the screen (p.10). As with most typographic advice, however, these suggestions are based upon years of experience coupled with personal preferences. They often do not represent a scientific formula for legible type in whole or in part. Hence the need for more empirical examinations.

Few have addressed the problems inherent to the rendering of type on a low-resolution, bitmapped CRT or LCD CRT. Up until 1995 when Microsoft Corporation commissioned Matthew Carter to design fonts for its Office 97 software suite, no one had attempted design typefaces specifically for on-screen viewing. These fonts are included in Office 97 as well as being freely downloadable for anyone's use from the WWW (Will-Harris 1998). The two common types of text used for on-screen viewing were originally designed for print output. They were never optimized for the CRT, however designers persist to use them due to the sparse alternatives. TrueType and PostScript were developed during the 1980s and were a boon to the desktop publishing revolution that occurred then. Back then the majority of type was intended to be output for hardcopy viewing. These fonts worked well because they vastly improved the imaging quality previously afforded by low-quality dot matrix output. However, as Will-Harris (1998) goes on to explain "the limitations inherent to reading on-screen especially the low-resolution of displays, mean that the digital designer has to be more careful about choosing typefaces that

are easily readable on-screen.” The problem with type designed for print being rendered on-screen lies in its basic letterforms and how they conform to the bitmap of the computer display.

Georgia and Verdana: Typefaces Optimized for the Computer Screen

To address the problem of conforming letterforms to the bitmaps of CRTs, Verdana and Georgia were designed specifically for on-screen viewing by Mathew Carter, co-founder of Bitstream, Inc. and a well-respected type designer with several typefaces designs to his credit including BellCentenial—a typeface designed specifically for use in printed telephone books. Tom Rinkner of Monotype did the hinting for these typefaces which were commissioned by Microsoft to optimize readability of both serif and sans serif fonts on-screen. Verdana, named for a region near Seattle, is a sans serif font similar in design to Arial. Georgia, named after a tabloid headline about alien heads found in the state of Georgia is a serif typeface resembling the serif typeface, Times (Will-Harris, 1998).

The design of these fonts demonstrated a "reversal of priorities" in that they were designed for the screen from the outset. Instead of working from outlines and then adjusting the bitmaps to represent them, a reverse approach was applied. The passage below illuminates this process:

"The process of instructing fonts by turning pixels on and off for each character at every point size is truly the art of a craftsman. Although the computer is used to adjust a letterform's outline at every size and resolution, it is a manual process in which the typographer decides what adjustments need to be made to render the best looking letter at every size. Each component of a letterform (serifs, stems, alignments, terminals) is given careful scrutiny to determine whether the resulting shape best represents the letter's outline. These types of decisions can only be made by a skilled typographer who is experienced with the rigors of low-resolution media. Fortunately, the time devoted to the font production process results in better quality and is very noticeable" (Monotype, 1998).

According to Will-Harris (1998) "Verdana does several things to maximize readability: its x-height is large, characters are extended (extra set width), increased letterspacing, bolds are optimized so as not to fill-in even at the four point setting, and curves are minimized in the counters. "Microsoft's Georgia is an impressive achievement in font design because it looks as sharp and clean on-screen as most type looks on-paper". In addition Georgia and Verdana are optimized for on-screen usage due to the fact that both have controlled letterspacing which ensures that letters never touch (Will-Harris 1998).

Will-Harris (1998b) notes that designers in the past have created new typefaces to take advantage of new advances in typesetting or printing medium. This need has never been more apparent than in today's burgeoning world of digital type. The fact that both Georgia and Verdana have been designed specifically with on-screen legibility in mind should prove to give it a decided advantage over other designs that only consider print. However, no research exists to investigate this contention.

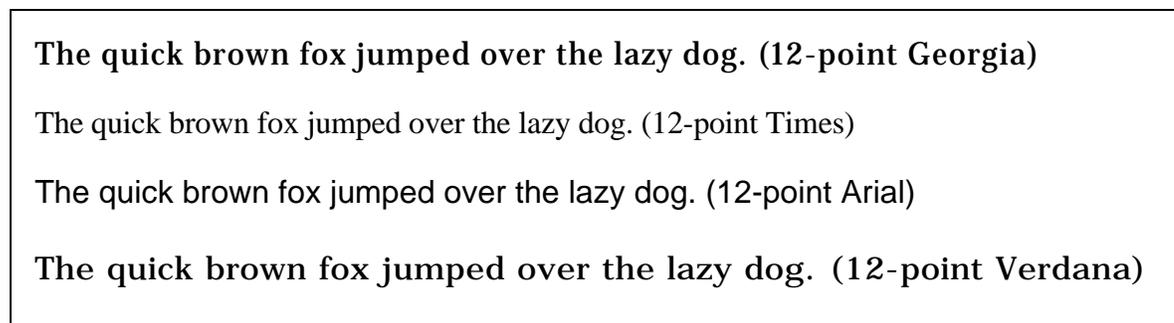


Figure 3: Georgia, Times, Arial, and Verdana

Typography on the World Wide Web

In 1997 recognized authority in WWW design Lynda Weinman wrote, “the web will never be a sophisticated visual environment until more font options are available” (p. 8). This quote illuminates the problem of the lack of design versatility for which the WWW is known. As opposed to print media— where the

designer is given control over the look of a publication up until it goes to press and is made permanent, the viewer of today's WWW documents commands control over many formatting variables. In many situations where electronic documents are viewed on-screen (e.g. the WWW), the reader is given control over the formatting options such as font, background color, and width of the document. It is reasonable to assume that the users would either modify the document to suit their own personal perceived preferences where these preference options are alterable by the browser software. For example, in *Netscape Navigator* (Netscape Corporation, 1998), a widely used interface (or browser) designed for the WWW, users have the option to control many formatting variables. This allows the browser to convert all hypertext according to the settings in the preferences of the browser. A user has the option of changing both the size and font of the hypertext displayed on the downloaded pages that they view. However, it could also be assumed that many readers would leave default formatting options in place for lack of concern.

The fact that the control of the typographic and layout design of electronic documents has been handed over to the discretion of the reader as opposed to the author(s) and designer(s) of the document (Gulick, 1996, p. 12), runs counter to the print-based paradigm of the past in which the author, the designer, or the publisher has had exclusive control over the format and typographic specifications of documents even up to the point that they go to press. Choices in typeface use prior to the advent of desktop publishing were

based upon either tradition of tried-and-true guidelines or for technical reasons. Today, however, our current system of networked communication-embodied in the WWW and enabled by the proliferation of desktop computers- allows anyone with the proper, inexpensive desktop computers and software to publish on the WWW. This has, in effect, spawned a revolution in the publishing of documents that are created with little concern for-or knowledge of-formerly recognized graphic design fundamentals. Therefore, many of the documents published on the WWW have eluded the traditions and guidelines of traditional printed materials. In studies by Tullis, Boynton, and Hersh's (1995) and Taylor (1990) subjects who were given the opportunity to rate typeface preferences, sans serif typefaces were preferred by the majority of the test subjects. However, in both studies, serif typefaces were shown to be more legible (i.e. able to be read faster). High (1997) supports this notion by stating: "The readability goals of the designer of the software now are the key to effective knowledge transfer, in fact, the content can be rendered useless all too easily by poor choices made by persons remote both in distance, objective and time."

In many cases WWW pages have no fixed (standardized) dimensions or typographic specifications. It is often left up to the reader to determine for example, the width of the window from which they are reading, or the font that they desire. Since WWW browsers give the user a readily available function with which to specify the font in which the hypertext will display, designers often

lose control of the typographic format of their electronic documents. This, in effect, makes the browser the typographer. Therefore, since it is virtually impossible for all readers to be skilled in or even aware of the effects of font choice on reading rate, it is conceivable that reading rate will ultimately suffer in some cases where readers make less than optimal font choices, not to mention the layout and design aspects of the page.

Changes are in the works, however, that will give an increased level of priority to authors of online documents. Innovations such as the tag, trued, OpenType, OpenFont, Cascading Style Sheets, and Microsoft Web Embedding Fonts Tool (WEFT) will give designers of on-screen text more freedom and authority over their documents. However, lack of consensus in this area is currently impeding a true standard.

The tag allows designers of HTML pages to specify fonts for the hypertext to display. However, this tag only works with fonts that have been previously installed in the fonts folder of the browser's computer (Microsoft, 1998). If that specified typeface is absent, then the browser defaults to another system font.

Microsoft Web Embedding Fonts Tool (WEFT) is another recent advance. WEFT is a free utility that lets you create 'font objects' that are linked to your WWW pages. "Font embedding for the WebFonts has been a feature of Microsoft applications such as Word and PowerPoint for several years. It allows the fonts used in the creation of a document to travel with that document,

ensuring that a user sees documents exactly as the designer intended them to be seen. Now, with the release of Microsoft Internet Explorer 4.0 for Windows, ... embedded fonts [are brought] to the WWW. This technology is set to change the look of the WWW, by empowering site designers to ensure their pages appear as they want them to" (Microsoft, 1998).

Microsoft's approach to embedded fonts within WWW pages is unique from previous methods in that the designers of this system "feel that font rasterization, font scaling, anti-aliasing, hinting, support and installation are issues best handled by the operating system and not by a browser based 'font displayer' ... individual glyph shapes should not be synthesized approximations of the original letterforms, but should remain true to the designer's intentions, and should be exactly the same as the glyphs contained in the original font... type designers' concerns need to be adequately addressed, and that font vendors' rights to produce 'no-embedding' fonts should be respected. According to Microsoft (1998), "to best achieve these ends (Microsoft) feels that the embedded font files should be based on proven TrueType and Acrobat Type 1 font technology, and are working closely with Acrobat Systems Inc. to implement an embedding solution based on the OpenType font file. Open Type allows fonts to contain either TrueType or PostScript data, or both. OpenType promises to make font development, management and use much easier." By attaching Cascading Style Sheets to structured documents on the WWW (e.g. HTML), authors and readers can influence the presentation of documents

without sacrificing device-independence or adding new HTML tags (W3C, 1997).

While these solutions are searching for a standardization resolution, a few simpler developments which get at the heart of the problem of on-screen text readability are screen-optimized fonts. These include typeface selections from three major vendors Microsoft, Acrobat, and Monotype. The lack of screen optimized typefaces can be accounted for by the following factors: (a) the difficulty of the WWW to transmit fonts (it is not as easy as transmitting pictures embedded in pages), (b) security, and (c) the complexity of the task (which requires an understanding of not only design, but technology and legibility factors (Will-Harris, 1998).

How Humans Read

What takes place during the human text interaction is still not entirely understood by scientists and researchers of the subject. However, the reading process is composed of the basic components of eye rhythm, fixations, eye span, thought units, configuration, recognition rate, familiarity, and reading rate (Rieck, 1997, 23-26).

Eye rhythm is the predisposition of readers in the Western world to scan a page of text line-by-line from left to right and top to bottom beginning in the upper-left corner of a page. There is a sweeping motion that the eye makes as it jumps from the end of one line to the beginning of the next. Fixations, or

"saccades", are momentary stopping points (approximately .25 seconds) that the eye makes in the midst of scanning a line of text in order to allow the brain to comprehend what it is visually processing. The legibility of a letter is known to decrease with its distance from the position where the eye is fixating. (Nazir, Jacobs, & O'Reagan, 1998; Anstis, 1974; Bouma, 1970,; Nazir, Heller, & Sussmann, 1992; Nazir, O'Reagan, & Jacob 1991; Olzak & Thomas, 1986; Townsend, Taylor, & Brown 1971). Therefore, reading distance beyond the point of fixation and perceived size of the text has a linear relationship.

Eye span refers to the radius around each point of fixation. An average eye-span radius extends about two inches around the point of fixation which loosely translates to about two and one-half words per fixation. Eye span is not an absolute measure, however. Readers' brains have the natural tendency to "chunk" or divide sentences into groups of words that express a cohesive idea according to the context of the sentence. For example, in the sentence "Christine kicked the ball," the two thought units are "Christine"—the person who acted—and "kicked the ball"—the action she performed.

Configuration refers to the distinctive shape patterns that written words create. Lowercase letters have a more distinctive shape than capital letters, therefore they can be perceived more quickly than uppercase letters. Because readers are frequently exposed to a word, they no longer have to "read" the word, but instantly recognize the meaning by the familiar shape of the group of letters. A common example is a "STOP" sign. The configuration of letters

composing words in a textual passage effect the recognition rate, or how quickly a person understands the words that he or she is reading. Often "familiarity" with the appearance of words—both the order of the letters and the typeface—effect the ease of reading. Naturally, distinctive patterns are more recognizable and memorable. Therefore, it is a widely held understanding that Roman or serif typefaces are more recognizable to readers. As a rule-of-thumb, serif typefaces are used to display body text (text set at 12-point or smaller). "Word recognition performance varies systematically as a function of where the eyes fixate in the word. Performance is maximal with the eye slightly left of the center of the word and decreases drastically to both sides of this optimal viewing position, or VPE" (Nazir, et al., 1998, p. 810).

Kennedy and Murray (1993, p. 251) state in their findings that "subjects read more slowly from a computer screen and the difference does not appear to be some trivial artifact of, for example, the orientation of the screen or the subject's posture." Instead they conjecture that the differences on reading rate arise from the distribution of the movements of the eyes of the reader. "Since normally, because of acuity limitations, the only information available in parafoveal vision concerns low-resolution features of letters; even when magnification provides better information, readers are unable to make use of it."

Nelson-Denny Reading Test

The *Nelson-Denny Reading Test* is a two-part, survey reading test that measures vocabulary development, comprehension, and reading rate. It is appropriate for testing high school and college students as well as adults. Part I is a fifteen-minute timed test which measures vocabulary. Part II is a twenty-minute test that measures reading comprehension and reading rate. The first sixty seconds of the comprehension test is used to determine reading rate. The documentation for the test suggests that the *Nelson-Denny Reading Test* can be administered in twenty minutes. This allows for time to distribute materials, complete the name and information grids, and provide directions. The most current edition of the *Nelson-Denny Reading Test* was published in 1993. It includes updated forms G and H. The emphasis has gradually shifted in the current versions of the *Nelson-Denny Reading Test* from reading rate measurement toward measurement of reading power (Riverside Publishing, 1998).

It should be noted that the passages used in the comprehension passages were "drawn from widely used, current high school and college texts" (Riverside Publishing, 1998). Furthermore, the authors of the current forms (Forms G and H) had the testing items screened for racial and gender bias. Items from previous versions of the test which were found to produce bias were removed, while the remaining items which were found to produce bias—but were determined to be essential to the test—were balanced in bias in

respect to the socio-economic categories which were taken into consideration (Murray-Ward, 1998).

CHAPTER THREE

Research Objectives

The intent of this research was to examine the readability of fonts designed for on-screen reading as compared with fonts designed for print output as measured by performance levels on a test of reading rate and reading comprehension. More specifically, this study was designed to empirically compare the reading performance of subjects who read text set in two different fonts (Georgia and Verdana) which were specifically designed for use on computer screens as compared with two fonts which were specifically designed for print output (Times and Arial).

This study was designed to test variations of the independent variables—font and presentation mode—on the dependent variables of reading rate and reading comprehension. Toward this end, an experiment was conducted in which subjects were tested for reading rate and reading comprehension using portions of a validated test instrument: the *Nelson-Denny Reading Test - Form G*. The tests were presented both on a computer screen (in one of two presentation modes—1-bit text and 8-bit text) and on paper as 600dpi laser printed output. Also, subjects were given versions of the test set in one of the fonts under examination (Georgia, Verdana, Times, or Arial). As a result, the entire sample selected for this study was divided into twelve

distinct treatment groups—each being tested for a different combination of font and/or presentation mode.

The four fonts examined in this experiment (Georgia, Verdana, Times, and Arial) are representative of widely distributed serif and sans serif typefaces, however, each pair is designed to function optimally on two different media. Georgia and Verdana were designed specifically to optimize readability of text appearing on a computer screen, whereas Times and Arial were designed specifically to enhance readability of text printed on paper. Each of these typefaces are widely distributed, standard fonts used for digital typography (Will-Harris, 1998).

Since it has been demonstrated in previous research that reading rate is dependent upon the interrelationship of both the peripheral vision process (visual acuity) as well as the cognitive processes of comprehension (Le Rohellec, Brettel, & Glassar, 1996), both dependent variables (reading rate and reading comprehension) were examined in this experiment. Statistical comparisons were made between screen-optimized fonts and digital fonts—designed specifically for text-on-paper output—as they appear both on-screen and on-paper. Four levels of the independent variable of font, Georgia, Times, Verdana, and Arial were used. Three levels of presentation mode were examined including (a) 1-bit rendering of on-screen text, (b) 600dpi rendering of text on-paper, and (c) 8-bit rendering of Adobe Portable Document format 8-bit text on a computer display.

The typographic variables selected for examination in this study were representative of four specific fonts (fonts are character sets which share the same *size* and *style*—see Appendix A - Typography Primer). The fonts, 12-point Times and 12-point Georgia, are representative of Roman typefaces commonly for the typesetting of body text (text set at 12-points or less) of documents. The fonts 12-point Arial and 12-point Verdana are representative of sans serif typefaces used most commonly for setting display text (for specifications of Roman and Sans Serif typefaces (see Appendix A).

These typographic (independent) variables in combination with the two (independent) variables of presentation mode (1-bit on-screen, 8-bit on-screen, and, 600dpi on paper) comprise the twelve treatment groups studied in this experiment (See *Table 1*).

Table 1

Treatment Groups

| | |
|-----------------|---|
| Group 1 | read 1-bit text set in the font Georgia on a computer display |
| Group 2 | read text rendered at 600dpi set in the font Georgia on paper |
| Group 3 | read 8-bit text set in the font Georgia on a computer display |
| Group 4 | read 1-bit text set in the font Times on a computer display |
| Group 5 | read text rendered at 600dpi set in the font Times on paper |
| Group 6 | read 8-bit text set in the font Times on a computer display |
| Group 7 | read 1-bit text set in the font Verdana on a computer display |
| Group 8 | read text rendered at 600dpi set in the font Verdana on paper |
| Group 9 | read 8-bit text set in the font Verdana on a computer display |
| Group 10 | read 1-bit text set in the font Arial on a computer display |
| Group 11 | read text rendered at 600dpi set in the font Arial on paper |
| Group 12 | read 8-bit text set in the font Arial on a computer display |

The dependent variables in this study were reading rate and reading comprehension. Results indicating significantly better scores in reading rate in any of the groups would substantiate or dispute evidence of the superiority in design of the screen optimized fonts "12-point Georgia" or "12-point Verdana" as they appear in any one of the three presentation modes investigated in this research.

Statistical Hypotheses

The following statistical hypotheses about reading rate and reading comprehension were tested using the data collected from the twelve treatment groups tested in this study:

1. There is no significant difference between the mean scores of reading rate across four levels of font variation (Times, Georgia, Arial, and Verdana) as they were displayed as 1-bit text on a computer screen.

$$H_0: \mu_{,1} = \mu_{,2} = \mu_{,3} = \mu_{,4}$$

$$H_a: \mu_{,1} \neq \mu_{,2} \neq \mu_{,3} \neq \mu_{,4}$$

2. There is no significant difference between the mean scores of reading comprehension across four levels of font variation (Times, Georgia, Arial, and Verdana) as they were displayed as 1-bit text on a computer screen.

$$H_0: \mu_{,1} = \mu_{,2} = \mu_{,3} = \mu_{,4}$$

$$H_a: \mu_{,1} \neq \mu_{,2} \neq \mu_{,3} \neq \mu_{,4}$$

3. There is no significant difference between the mean scores of reading rate across four levels of font variation (Times, Georgia, Arial, and Verdana) as they were displayed as 8-bit text on a computer screen.

$$H_0: \mu_{.5} = \mu_{.6} = \mu_{.7} = \mu_{.8}$$

$$H_a: \mu_{.5} \neq \mu_{.6} \neq \mu_{.7} \neq \mu_{.8}$$

4. There is no significant difference between the mean scores of reading comprehension across four levels of font variation (Times, Georgia, Arial, and Verdana) as they were displayed as 8-bit text on a computer screen.

$$H_0: \mu_{.5} = \mu_{.6} = \mu_{.7} = \mu_{.8}$$

$$H_a: \mu_{.5} \neq \mu_{.6} \neq \mu_{.7} \neq \mu_{.8}$$

5. There is no significant difference between the mean scores of reading rate across four levels of font variation (Times, Georgia, Arial, and Verdana) as they were displayed as 600dpi text on paper.

$$H_0: \mu_{.9} = \mu_{.10} = \mu_{.11} = \mu_{.12}$$

$$H_a: \mu_{.9} \neq \mu_{.10} \neq \mu_{.11} \neq \mu_{.12}$$

6. There is no significant difference between the mean scores of reading comprehension across four levels of font variation (Times, Georgia, Arial, and Verdana) as they were displayed as 600dpi text on paper.

$$H_0: \mu_{,9} = \mu_{,10} = \mu_{,11} = \mu_{,12}$$

$$H_a: \mu_{,9} \neq \mu_{,10} \neq \mu_{,11} \neq \mu_{,12}$$

Research Design

This research study implemented a quasi-experimental design to test the effects of the independent variables (variables—font and presentation mode) upon the dependent variables (reading rate and reading comprehension). The variables in this study are referred to as categorical variables. That is, the variables at hand consist of four different categories of text embodied by four different fonts and three different types of presentation mode each imparting distinctive physical characteristics to the body of text. A "categorical independent variable post-measure only" design is useful when comparing the effects of a set of categorical variables. It was suggested that this type of design is applicable when the independent variables are representative of any number of different categories. Therefore it was implemented into the study which includes twelve different categories (See *Table 1*).

It must also be pointed out that control in this type of design is attained by random assignment of subjects to the different categories of independent variables which, in effect, equates the groups in all respects except for the effects imposed by the independent variables. Therefore, it can be assumed that any differences in the means of the groups (categories of text) tested in this study derived from the treatment introduced to the particular treatment group (category) to which they were randomly assigned. This alleviates any need to include control groups in the research design (Pedhazur & Schmelkin, 1991, pp. 268-69).

Table 2

*Factorial Design of the Study Illustrating the Categorical Independent Variable
Post-Measure Only Design*

| | | |
|----------------|-----|---------------------------------|
| R ₁ | --> | O _R , O _C |
| R ₂ | --> | O _R , O _C |
| R ₃ | --> | O _R , O _C |
| R ₄ | --> | O _R , O _C |
| R ₅ | --> | O _R , O _C |
| R ₆ | --> | O _R , O _C |
| R ₇ | --> | O _R , O _C |

$R_8 \rightarrow O_R, O_C$

$R_9 \rightarrow O_R, O_C$

$R_{10} \rightarrow O_R, O_C$

$R_{11} \rightarrow O_R, O_C$

$R_{12} \rightarrow O_R, O_C$

Where:

R_1 = Read 1-bit text set in the font “Times” on a computer display

R_2 = read 1-bit text set in the font “Georgia” on a computer display

R_3 = read 1-bit text set in the font “Arial” on a computer display

R_4 = read 1-bit text set in the font “Verdana” on a computer display

R_5 = read 8-bit text set in the font “Times” on a computer display

R_6 = read 8-bit text set in the font “Georgia” on a computer display

R_7 = read 8-bit text set in the font “Arial” on a computer display

R_8 = read 8-bit text set in the font “Verdana” on a computer display

R_9 = read text rendered at 600dpi set in the font “Times” on paper

R_{10} = read text rendered at 600dpi set in the font “Georgia” on

paper

R_{11} = read text rendered at 600dpi set in the font “Arial” on paper

R_{12} = read text rendered at 600dpi set in the font “Verdana” on
paper

O_R = Observed randomized group mean scores for reading rate

O_C = Observed randomized group mean scores for reading
comprehension

Table 3

Factorial Design 3 X 4 Analysis of Variance (ANOVA)

| | Presentation Mode | | |
|---------|-------------------|-----------------|-----------------|
| | 1-bit on screen | 8-bit on screen | 600dpi on paper |
| Times | Group 1 [n=22] | Group 5 [n=22] | Group 9 [n=22] |
| Georgia | Group 2 [n=22] | Group 6 [n=22] | Group 10 [n=22] |
| Arial | Group 3 [n=22] | Group 7 [n=22] | Group 11 [n=22] |
| Verdana | Group 4 [n=22] | Group 8 [n=22] | Group 12 [n=22] |

Population and Sample

The sample selected for this study was composed of students at Illinois State University. These subjects were selected from the class rosters of courses offered in the Department of Industrial Technology during the Spring semester of 1999 at Illinois State University in Normal, Illinois. See Appendix C for a listing of the titles of the courses from which subjects for this test were sampled. In order to gain as precise a measurement as possible, it was determined that a sample size of 264 subjects was to be used as a pool for collecting data in this study. This number allows for 22 subjects to be randomly assigned to each of twelve treatment groups which, for data analysis purposes, correspond to twelve different cells of the research design matrix [See Table 2]. According to Levin and Kaplan (1972) this provides an adequate sample size for the number of treatment groups in this study.

Over 95 percent of the subjects were Industrial Technology majors. Industrial Technology is one of several departments in the College of Applied Science and Technology at Illinois State University in Normal, Illinois. Other majors that participated, and accounted for less than five percent of the population included Applied Computer Science, Agricultural Science, Family and Consumer Science, and Military Science. All of these majors are grouped into the College of Applied Science and Technology. The College of Applied Science and Technology (CAST) is composed of over 3325 majors. The College contains eight diverse departments including: Agriculture; Applied

Computer Science; Criminal Justice Sciences; Health, Physical Education and Recreation; Health Sciences; Family and Consumer Sciences; Industrial Technology; and Military Science (Illinois State University, 1999). Classes selected for the study ranged from introductory-level to upper-level.

The median age of the subjects was 21.5 years. Males constituted 74 percent of the sample. The ethnic make-up of the sample was 84 percent White Non-Hispanic, 8 percent Black Non-Hispanic, 5 percent Asian or Pacific Islander, and 3 percent Hispanic.

As Holmes (1986) pointed out, readability research has traditionally concluded that a combination of reader and typographic variables affect reading rate and reading comprehension. While no pretense can be made that would suggest that this sample is representative of universal reading ability, this group could be described as generally representative of college students of average ability who possess a predisposition toward technology and who are currently enrolled at a Midwestern university. Therefore, it represents a homogenous sample in and of itself and cannot be truly generalized to the entire population of the world's readers. However, due to the homogeneity of subject sampled, the results of this research could be "hypothetically generalized" to portions of the population who read at or near the college-level and share similar demographic characteristics.

Test Instrument

To a large extent, measures of reading rate and reading comprehension have been cited in research as determinants of readability (Tullis et. al., 1995; Turner, 1982; Taylor, 1990; Lough, 1982; Journa, 1989; Holmes, 1986; Canary, 1983). Therefore, this research followed a similar methodology as other previous studies of this nature by employing a previously validated test for measuring reading rate and reading comprehension: the *Nelson-Denny Reading Test* - Form G.

The *Nelson-Denny Reading Test* - Form G has been validated and represented most frequently throughout similar experimental research on measures of the rate and comprehension of reading. Although it has had its critics throughout the years since its origination in 1929 (Murray-Ward, 1998; Smith, D. K., 1998), it has been widely used and remained generally unchanged since then. Therefore, it was deemed appropriate for the purposes of assessing both reading rate and reading comprehension of the subjects in this experiment. Only the reading rate and reading comprehension portions of this test were used for the purposes of this study.

All subjects were given the same textual passages to read. Although the content of each test was identical, the fonts with which the text was displayed varied with each treatment group. The hard copy versions of the test were reproduced with the exact same typographic specifications as the original test. The 1-bit on-screen versions of the test were reproduced using Microsoft Word

with the identical typographic specifications as the original test. The 8-bit on-screen version of the test was reproduced with the exact same typographic specifications as the original and then was converted to Adobe Portable Document format using Adobe Acrobat. To ensure that the fonts displayed at the same physical size in all cases, subjects taking the test using the test form displayed in 1-bit on-screen text were directed to set the "Zoom" setting in Microsoft Word to 90 percent, while those subjects using the test form displayed in 8-bit on-screen text were directed to set the "Zoom to..." setting in Adobe Acrobat Reader 3.0 to 125 percent.

Testing Procedures

Prior to the administration of the treatment, subjects were screened via a written questionnaire [See Appendix D] to determine if their vision was normal, corrected, or impaired. Those subjects whose vision was impaired for any reason were excluded from the sample. Despite the fact that their data would not be used for statistical analysis, those subjects who indicated that their vision was impaired were not discouraged from taking the test.

Additionally, demographic information was gathered prior to the administration of the treatment via a simple questionnaire [see Appendix D] which asked subjects to proffer personal information regarding gender, age, ethnicity, and native tongue. It was determined that the data gathered in this

demographic questionnaire would be held indefinitely for undetermined future analytical purposes.

Subjects were selected from class rosters of courses offered in the Department of Industrial Technology during the spring semester of 1999 at Illinois State. Subjects were assigned to treatment groups by means of simple random sampling (SRS). As prescribed by Babbie (1990, pp. 83-84) each student name was numbered according to the alphabetical order in which they appear on the class roster. Then treatment groups were assigned using systematic random assignment (SRA). This method was also used in a similar study conducted by Gerig, Nibbelink, & Hoover (1992) which also employed standardized tests as the basis for comparison of typographic formatting variables.

Federal regulations and Illinois State University policy require that all research involving humans as subjects be approved by the Illinois State University Institutional Review Board (IRB). Permission to test human subjects (Illinois State University students) was sought and granted prior to testing and final written approval was received from James C. Palmer, Chairperson of the Illinois State University IRB. The IRB classified this study as "exempt" from the Board's oversight.

Permission was also granted by instructors of the classes from which the sample was extracted as well as from the administration of the Illinois State University Department of Industrial Technology. Actual administration and

scheduling of reading rate tests was accomplished through assistance from graduate assistants and undergraduate student assistants. Subjects were informed of the intent of the research and asked if they were willing to participate.

To measure reading rate, subjects participating in the study were asked to read a short passage of text from the *Nelson-Denny Reading Test*. Instructions to the test were delivered and time was kept by the researcher. All passages were identical in content and format to the Reading Rate and Reading Comprehension portions of *Nelson-Denny Reading Test - Form G*. Students taking the test from computer screens used a computer display set to a pixel resolution of 1280 x 1024. Subjects who took the test via a computerized test form were instructed not to use the computer mouse for scrolling purposes. Subjects were instructed to use only the "PageUp" and "PageDown" keys to navigate the test form. This eliminated the scrolling variable which would potentially create inconsistency between the subject reading from hard copy and the subjects reading from computer screens.

In accordance with the directions to the test, this portion of the test was timed for one minute. At the end of one minute, all subjects directed to indicate the last word they had read on the answer sheet provided by the test administrator. Students were also instructed to immediately continue with the remainder of the test as soon as they had marked their one-minute stopping point on the answer sheet. After twenty minutes had elapsed, time was called

by the test administrator. Subjects were instructed to stop working on the test at this point.

A questionnaire followed the actual testing process. The questionnaire was intended to extract subjective information about the readers' typeface preferences and the overall impression of the legibility of the text. Specific questions addressed the ability of the reader to distinguish between the "e's" from the "o's", etc.

Administration of the Test

In order to guard the anonymity of the subjects and/or the confidentiality of their test scores, no personal identification information was requested or gathered [see Appendix D]. Subjects were advised not to write their names or make any distinguishing marks on any of the testing materials. However, as sample subjects were taken from class groups and randomly assigned to treatment groups directly from class rosters, names of subjects become readily available. In order to ensure complete confidentiality, all class rosters used in the assignment of subjects were destroyed (shredded) after use. Furthermore, no report of class titles was mentioned in the accompanying research report. Also, any demographic information collected prior to the study will not require any form of identification information. Demographic information requested dealt with gender of the subject, quality/condition of the subjects' eyesight, and the subjects' native language [see Appendix D].

Actual test forms and demographic information questionnaires were kept on file in the office of Assistant Professor, Eric Weisenmiller (Nelson Smith Building Room 3) until the project was completed (July, 1999). Access to this room is restricted to Eric Weisenmiller and custodial employees of the Nelson Smith Building.

Subjects were informed of the nature of their participation in the project prior to the administration of the test both orally by the administrator of the test. They were also informed of the strictly voluntary nature of their participation by their instructor prior to entering the testing facility (Turner Hall Room 206). Confidentiality of the subjects' participation in the test and the data gathered during testing were assured both orally and in written form.

Using an adaptation of the *Nelson-Denny* scoring procedure, the number of words read by each subject were determined and grouped by comparing the stopping point of each subjects' test form with a master list. A mean score for each group was calculated. Then the mean scores in each cell were used to run statistical analyses.

Variables

The dependent variables in this study were reading rate and reading comprehension. The independent variables included four typefaces in which text for the reading test was set. The two serif typefaces included Georgia (a typeface designed specifically for on-screen viewing), and "Times" (a typeface

designed for digital typesetting and text-on-paper output). The two sans serif typefaces included Verdana (a typeface designed specifically for on-screen viewing), and Arial (a typeface designed for digital typesetting and text-on-paper output). All typefaces were set at 12-point and were read either on a CRT or on 600dpi laser printed on white paper. All text was set to a standard 12-point size. This is generally recognized as the most common default point size in applications which display text. In addition, the majority of typeface designers construct their typefaces to perform optimally at standard or "critical point sizes" (Carter, quoted in Will-Harris, 1998)—these being 9 point, 10-point, and 12-point.

Control of Extraneous Variables

In a study of this nature, in which tangible variables under consideration are potentially very numerous, it is essential to provide an explanation, not only of the variables being examined, but also of those chosen not to be included for examination. The goal of this research was not to solve all problems pertaining to on-screen text legibility, but rather to examine two typographic variables that potentially effect readability—namely font and presentation mode. Therefore, this study held constant the environmental variables of noise and lighting, the physiological variables of eye fatigue and reading distance, and hardware variables such as display luminance, chromaticity, resolution, brightness, and

contrast. Often these variables, or combinations of these variables, vary widely due to personal preference of a specific reader.

As noted in (Le Rohellec, et. al., 1996) contrast polarity has little effect on performance although black on white characters gained generally better comfort ratings [as contrast increased]. Therefore contrast was standardized to a medium level on all computer displays for consistency's sake. In order to control for extraneous variability in the quality of the appearance of the image of the text as it appears on the display, all displays were of the same make and model. Along with brightness and chromaticity, contrast was set to the default setting of the 17" Megaplus[®] computer display. All displays were reset to their default settings prior to the administration of the tests.

It has been stated in Journa (1995) that display resolution (image quality) is a function of reading rate. In other words, as image quality increases, reading rate increases. This was true regardless of medium—CRT or print. And, as variables associated with display quality are numerous (dot pitch, luminance, chromaticity, aspect ratio, refresh rate, etc.) and innovations are ongoing, the researcher has chosen to hold all display variables constant in order to concentrate exclusively upon the typographic variables at work as they exist. The existing limitation to on-screen typography being its lack of resolution, all text examined in this study—both on-screen and text-on-paper—were displayed at a resolution of 100 ppi.

Although there may be reason to infer that lighting conditions may have subtle effects on the primary variables under examination in this study. Although, lighting effects on reading performance may be an appropriate avenue of research, those effects are not of primary concern in this study. Therefore, as prescribed in the *Nelson-Denny Reading Test* "Preparation for Testing" specifications, the room should be "well lighted" (Brown, Fishco, & Hanna, 1993). All lighting was consistent throughout all testing sessions. All exterior light was blocked from entering through four small 4' x 3' windows located in the corners of the computer lab used in this study.

Statistical Procedures and Data Analysis

Mean scores were gathered from each of the twelve groups of twenty-two subjects receiving a different treatment. Mean scores from all eight groups were analyzed and compared against the error term using a 3 X 4 two-way multi-factor analysis of variance (ANOVA) as prescribed by Pedhazur (1991) in order to discover any significance among any of the eight treatment groups. (p.269) This test will indicate both the main effects of the independent variables as well as any interactions among the independent variables (e.g. the effect of media on a particular font's legibility). The statistical analyses were run using Minitab™ statistical software to determine whether the F-ratios among the group means are significant at the 0.05 level of significance.

CHAPTER FOUR

Introduction

The purpose of this study was to determine empirically whether the typefaces Georgia and Verdana (which have been optimized for on-screen viewing) significantly improve reading performance as measured by reading rate and/or reading comprehension as compared to Times and Arial—both of which are digital fonts designed specifically for text-on-paper output. In addition, all typefaces were tested as they appear both on a computer display and on paper.

In order to measure reading rate and reading comprehension, an experiment was conducted in which college students were tested for reading rate and reading comprehension the *Nelson-Denny Reading Test - Form G*. The test was administered either on a computer screen or on paper—depending on which treatment group the subject was assigned. Also, subjects were given versions of the test typeset in any one of the typefaces under examination (Georgia, Verdana, Times, or Arial). As a result, the entire sample selected for this study was divided into twelve distinct treatment groups—each being tested for a different combination of typeface and/or media type.

After gathering both reading rate and reading comprehension data from the test scores of the participants, statistical comparisons were made between

the overall scores of those subjects who took the tests that were displayed using both screen-optimized typefaces and digital typefaces in combination with the two different presentation modes—on paper and on-screen. Four levels of the independent variable of font, Georgia, Times, Verdana, and Arial were used. Three levels of display were examined including (a) 1-bit rendering of on-screen text; (b) 8-bit rendering of Adobe Portable Document format grayscale text on a computer; and (c) display 600dpi rendering of text on paper. Thus, the sample was divided into twelve distinct treatment groups.

Analysis of Data

The purpose of this data analysis was to:

1. Distinguish differences in reading rate among or between twelve different levels of typeface variation and presentation mode.
2. Distinguish differences in reading comprehension among or between twelve different levels of typeface variation and presentation mode.
3. Distinguish any significant interactions between the two independent variables of typeface and presentation mode across four levels of typeface variation and three levels of presentation mode.

The data collection produced a total of 264 usable instruments that could be analyzed for both reading rate and reading comprehension. A balanced, or orthogonal, approach was employed to analyze the groups due to their equality of size (each group contained 22 members). The 264 subjects

were divided into twelve different sample groups each groups having been given one of twelve different versions of the test instrument. Each subject yielded two raw scores—one representing reading rate, the other reading comprehension. Collectively, each sample group produced a sample group mean score for both reading rate and reading comprehension. Those group means are listed in Table 4 and Table 5.

Table 4

Group Means (Standard Deviations) for Reading Rate Test

| Typeface | 1-bit On-screen | 8-bit On-screen | On-paper |
|----------|-----------------|-----------------|----------------|
| Times | 198.09 (49.26) | 219.14 (84.91) | 206.18 (47.28) |
| Georgia | 213.59 (81.22) | 207.55 (53.18) | 201.86 (75.23) |
| Arial | 188.50 (53.80) | 239.82 (61.76) | 190.45 (53.90) |
| Verdana | 187.82 (59.02) | 221.23 (66.70) | 206.05 (36.03) |

Note. Maximum possible score = 601.

Table 5

Group Means (Standard Deviations) for Reading Comprehension Test

| Typeface | 1-bit On-screen | 8-bit On-screen | On-paper |
|----------|-----------------|-----------------|----------------|
| Times | 24.727 (6.025) | 27.091 (5.554) | 27.273 (5.513) |
| Georgia | 23.000 (8.118) | 24.818 (7.817) | 27.000 (7.171) |
| Arial | 23.818 (7.055) | 28.591 (5.509) | 25.136 (7.344) |
| Verdana | 23.273 (5.775) | 28.000 (6.539) | 27.455 (6.501) |

Note. Maximum possible score = 36.

The group means for both reading rate and reading comprehension were analyzed using two two-way analysis of variance (ANOVA) tests. This test afforded the capability of determining significant differences among the sample group means. The *F*-values were tested for statistical significance at the .05 level.

Group Homogeneity

Among the assumptions that the test of ANOVA makes concerning its population parameters are homogeneity of variance, normality, independence observations, and the null hypothesis (Howell, 1997, 302-303). To ensure homogeneity of variance of the groups, an *F*-Max test was conducted. The *F*-Max test indicated that the variances of the sample groups were not significantly different. The critical value of 3.76 was found for comparison from the *F*-Max table corresponding to the degrees of freedom derived from the

sample groups and the number of subjects in each group [12, 22]. A value of 2.357 was derived from the reading rate variance scores by subtracting the largest from the smallest. Using the identical method, a value of 1.474 was obtained for the reading comprehension scores. Since both values did not exceed the critical value of $F_{cv} (12, 22) p < .05 = 3.76$, homogeneity of variance established. Therefore, the ANOVA could be conducted.

Testing the Null Hypotheses

The following sections describe the comparisons of the findings of the data analysis with the statistical hypotheses on reading rate and reading comprehension. In order to test for significance the following statistical operations and/or comparisons were employed:

1. Testing for significance in the interaction of presentation mode and font.
2. Testing for significance of each main effect (font or presentation mode on reading rate or reading comprehension).
3. Isolating significant difference between main effects using pairwise comparison.

Reading Rates

A two-way ANOVA was used to test the data gathered from the reading rate portion of the reformatted *Nelson-Denny Reading Test*. Since each cell of sample group means was equal in size, a balanced two-way ANOVA was used to analyze the test scores for reading rate. The results of the two-way ANOVA indicated the following:

1. There was no overall significant difference between the reading rates of sample groups who read text from the four different levels of variation in font ($F = 0.03$, $df = 3$, 252) at the 0.05 level of significance.
2. There was no overall significant interaction between font and presentation mode in respect to reading rate ($F = 1.07$, $df = 6$, 252) at the 0.05 level of significance.
3. There was, however, significant difference among the reading rates of sample groups who read text from the three different levels of variation in presentation mode ($F = 4.12$, $df = 2$, 252) at the 0.05 level of significance.

See Table 6 for a summary table of the ANOVA for reading rate.

It should be noted that since the reading test administered in this research was timed, the dependent variable of reading comprehension was in some cases affected by a subjects reading rate. Due to the fact that—in some cases—lack reading rate prohibited subjects from finishing the test of comprehension resulting in incomplete (lower) comprehensions scores.

Therefore, the results of this test should be tempered with the fact that reading comprehension is a concomitant variable of reading rate.

Table 7

Summary of Two-way ANOVA of Reading Rate Scores

| Source | <i>df</i> | SS | MS | <i>F</i> | <i>p</i> |
|--------------|-----------|---------|----------|----------|----------|
| Typeface (A) | 3 | 339 | 1130.039 | 0.03 | 0.993 |
| Mode (B) | 2 | 31421 | 15710 | 4.12 | 0.017 |
| A x B | 6 | 24574 | 4096 | 1.07 | 0.379 |
| Error | 252 | 961827 | 3817 | | |
| Total | 263 | 1018161 | | | |

= .05

Table 8

One-Way ANOVA for Reading Rate Across Three Levels of Presentation Mode

| Source | <i>df</i> | SS | MS | <i>F</i> | <i>p</i> |
|--------|-----------|---------|-------|----------|----------|
| Mode | 2 | 1421 | 15710 | 4.16 * | 0.017 |
| Error | 261 | 986740 | 3781 | | |
| Total | 263 | 1018161 | | | |

* $p < .05$ [$F_{.05}(2, 261) = 3.03$]

Table 9

Descriptive Statistics for Reading Rate Scores across Three Levels of Presentation Mode

| Level | <i>n</i> | <i>X</i> | <i>s</i> |
|-----------------|----------|----------|----------|
| 1-bit On-screen | 88 | 197.00 | 61.86 |
| 8-bit On-screen | 88 | 221.93 | 67.47 |
| 600dpi On-paper | 88 | 201.14 | 54.42 |

Table 10

Tukey's Pairwise Comparisons for Reading Rate Scores Across Three Levels of Presentation Mode

| | 1-bit On-screen | 8-bit On-screen |
|-----------------|-----------------|-----------------|
| 8-bit On-screen | -46.6* | |
| | -3.2 | |
| 600dpi On-paper | -25.8* | -0.9 |
| | 17.6 | 42.5 |

* $p < .05$ [$F_{.05}(2, 261) = 3.31$]

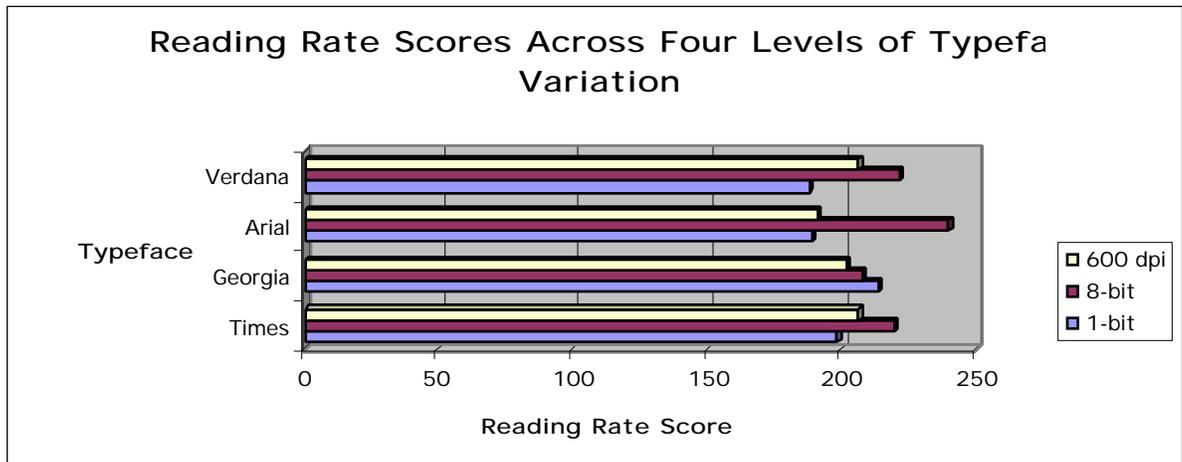


Figure 4. The results of the reading rate test across the different levels of typeface with respect to the mode of presentation.

Reading Comprehension

A two-way ANOVA was used to test the data gathered from the reading comprehension portion of the reformatted *Nelson-Denny Reading Test*. Since each cell of sample group means was equal in size, a balanced two-way ANOVA was also used to analyze the test scores for reading comprehension.

The results of the two-way ANOVA indicated the following:

1. There was no overall significant difference between the reading comprehension scores of sample groups who read text from the four different levels of variation in font ($F = 0.62$, $df = 3, 252$) at the 0.05 level of significance.

2. There was no overall significant interaction among font and presentation mode with respect to reading comprehension ($F = 0.81, df = 6, 252$) at the 0.05 level of significance.
3. There was, however, statistical significance found between the reading comprehension scores of sample groups who read text from the three different levels of variation in presentation mode ($F = 6.97, df = 2, 252$) at the 0.05 level of significance. See *Table 7* for a summary table of the ANOVA for reading comprehension.

Table 11

Summary of Two-way ANOVA of Reading Comprehension Scores for Reading Comprehension

| Source | | <i>df</i> | <i>SS</i> | <i>MS</i> | <i>F_p</i> |
|--------------|-----|-----------|-----------|-----------|----------------------|
| Typeface (A) | 3 | 82.30 | 27.43 | 0.62 | 0.601 |
| Mode (B) | 2 | 614.10 | 307.05 | 6.97 | 0.001 |
| A x B | 6 | 212.72 | 35.45 | 0.81 | 0.567 |
| Error | 252 | 11096.82 | 44.03 | | |
| Total | 263 | 12005.94 | | | |

= .05

Table 12

One-Way ANOVA on Comprehension Scores across Three Levels of Presentation Mode

| Source | <i>df</i> | SS | MS | <i>F</i> | <i>p</i> |
|--------|-----------|---------|-------|----------|----------|
| Mode | 2 | 614.1 | 307.0 | 7.03* | 0.001 |
| Error | 261 | 11391.8 | 43.6 | | |
| Total | 263 | 12005.9 | | | |

* $p < .05$ [$F_{.05}(2, 261) = 3.03$]

Table 13

Descriptive Statistics of Comprehension Scores across Three Levels of Presentation Mode

| Level | <i>n</i> | <i>M</i> | <i>s</i> |
|-----------------|----------|----------|----------|
| 1-bit On screen | 88 | 23.705 | 6.721 |
| 8-bit On-screen | 88 | 27.125 | 6.475 |
| 600dpi On-paper | 88 | 26.716 | 6.621 |

Table 14

Tukey's Pairwise Comparisons of Comprehension Scores across Three Levels of Presentation Mode

| | 1-bit On-screen | 8-bit On-screen |
|-----------------|-----------------|-----------------|
| 8-bit On-screen | -5.752* | |
| | -1.089 | |
| 600dpi On-paper | -5.342* | -1.922 |
| | -0.680 | 2.740 |

* $p < .05$ [$F_{.05}(2, 261) = 3.31$]

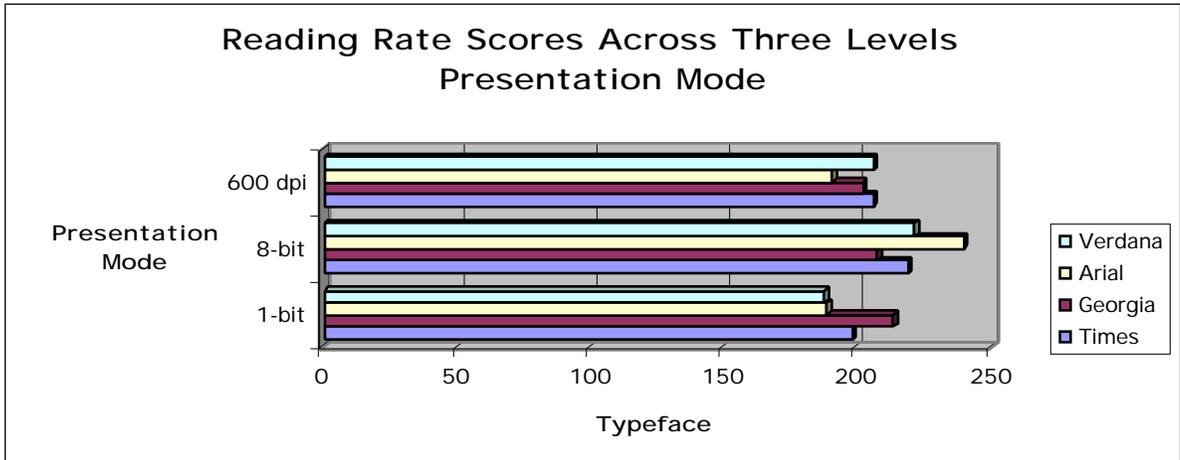


Figure 5. The results of the reading comprehension test across the different modes of presentation with respect to the four levels of typeface variation.

Findings

Hypothesis 1: There was no significant difference between the mean scores of reading rate across four levels of typeface variation (Times, Georgia, Arial, and Verdana) as they were displayed as 1-bit text on a computer screen.

A one-way ANOVA on rate across all four levels of typeface variation as presented as 1-bit text on a computer screen revealed no significant difference between scores of the two sample groups ($F = 0.82$, $df = 3, 84$) at the 0.05 level of significance. Therefore, hypothesis 1 could not be rejected.

Table 15

*One-Way ANOVA of Reading Rate Scores Across Four Levels of Typeface
Variation as They Were Displayed as 1-bit Text on a Computer Screen*

| Source | <i>df</i> | SS | MS | <i>F</i> | <i>p</i> |
|--------|-----------|--------|------|----------|----------|
| Form | 3 | 9526 | 3175 | 0.82 | 0.484 |
| Error | 84 | 323424 | 3850 | | |
| Total | 87 | 332950 | | | |

=.05

Table 16

*Descriptive Statistics of Reading Rate Scores across Four Levels of Typeface
Variation as They Were Displayed as 1-bit Text on a Computer Screen*

| Level | <i>n</i> | <i>X</i> | <i>s</i> |
|---------|----------|----------|----------|
| Times | 22 | 198.09 | 49.26 |
| Georgia | 22 | 213.59 | 81.22 |
| Arial | 22 | 188.50 | 53.80 |
| Verdana | 22 | 187.82 | 59.02 |

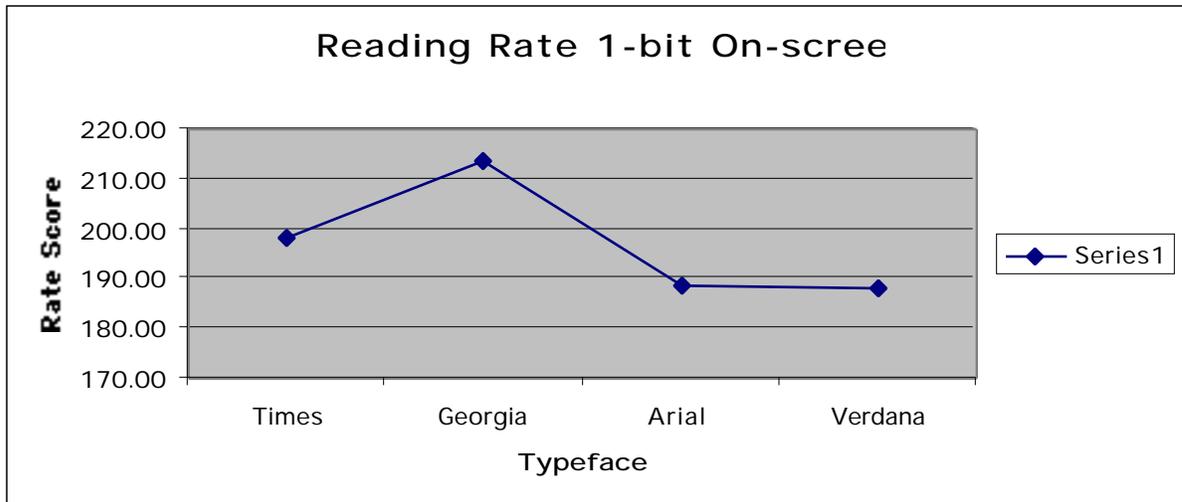


Figure 6. The results of the reading rate test across the four different levels of typeface as they were presented using 1-bit text on a computer screen.

Hypothesis 2: There was no significant difference between the mean scores of reading comprehension across four levels of typeface variation (Times, Georgia, Arial, and Verdana) as they were displayed as 1-bit text on a computer screen.

A one-way ANOVA on comprehension across all four levels of typeface variation as presented as 1-bit text on a computer screen revealed no significant difference between scores of the two sample groups ($F = 0.28$, $df = 3, 84$) at the 0.05 level of significance. Therefore, hypothesis 2 could not be rejected.

Table 17

ANOVA of Reading Comprehension Scores across four levels of typeface variation as displayed as 1-bit text on a computer screen

| <i>Source</i> | <i>df</i> | <i>SS</i> | <i>MS</i> | <i>F</i> | <i>p</i> |
|---------------|-----------|-----------|-----------|----------|----------|
| Form | 3 | 38.3 | 12.8 | 0.28 | 0.843 |
| Error | 84 | 3892.0 | 46.3 | | |
| Total | 87 | 3930.3 | | | |

=.05

Table 18

Descriptive Statistics of Reading Comprehension Scores across Four Levels of Typeface Variation as Displayed As 1-bit Text on A Computer Screen

| Level | <i>n</i> | <i>X</i> | <i>s</i> |
|---------|----------|----------|----------|
| Times | 22 | 24.727 | 6.025 |
| Georgia | 22 | 23.000 | 8.118 |
| Arial | 22 | 23.818 | 7.055 |
| Verdana | 22 | 23.273 | 5.775 |

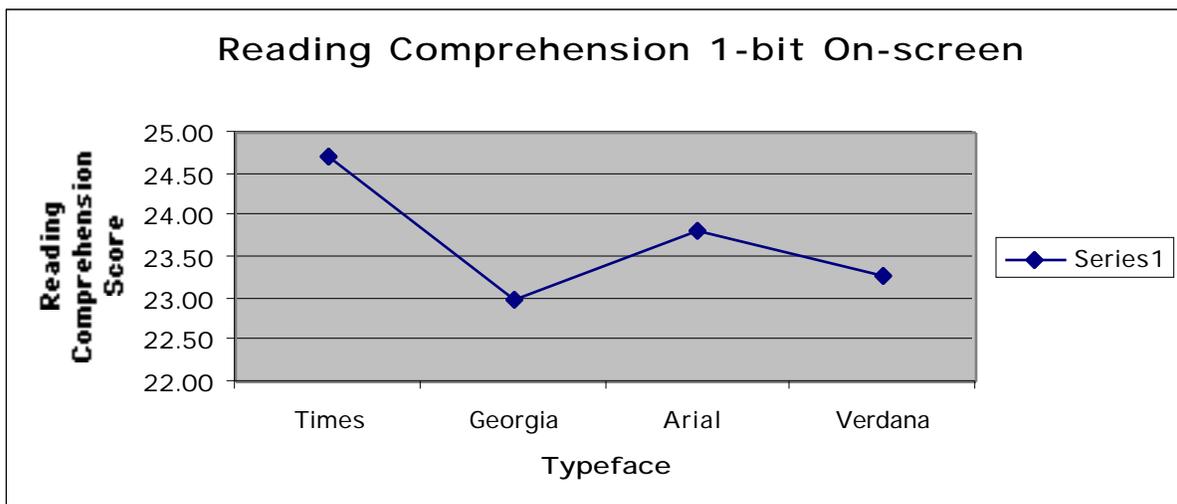


Figure 7. The results of the reading comprehension test across the four different levels of typeface as they were presented using 1-bit text on a computer screen.

Hypothesis 3: There was no significant difference between the mean scores of reading rate across four levels of typeface variation (Times, Georgia, Arial, and Verdana) as they were displayed as 8-bit text on a computer screen.

A one-way ANOVA on rate across all four levels of typeface variation as presented as 8-bit text on a computer screen revealed no significant difference between scores of the two sample groups ($F = 0.86$, $df = 3, 84$) at the 0.05 level of significance. Therefore, hypothesis 3 could not be rejected.

Table 19

One-Way ANOVA of Reading Rate Scores across Four Levels of Typeface Variation as They Were Displayed as 8-bit Text on a Computer Screen

| <i>Source</i> | <i>df</i> | <i>SS</i> | <i>MS</i> | <i>F</i> | <i>p</i> |
|---------------|-----------|-----------|-----------|----------|----------|
| Form | 3 | 11774 | 3925 | 0.86 | 0.466 |
| Error | 84 | 384325 | 4575 | | |
| Total | 87 | 396100 | | | |

=.05

Table 20

Descriptive Statistics of Reading Rate Scores across Four Levels of Typeface Variation as They Were Displayed as 8-bit Text on a Computer Screen

| Level | <i>n</i> | <i>X</i> | <i>s</i> |
|---------|----------|----------|----------|
| Times | 22 | 219.14 | 84.91 |
| Georgia | 22 | 207.55 | 53.18 |
| Arial | 22 | 239.82 | 61.76 |
| Verdana | 22 | 221.23 | 66.70 |

Hypothesis 4: There was no significant difference between the mean scores of reading comprehension across four levels of typeface variation (Times, Georgia, Arial, and Verdana) as they were displayed as 8-bit text on a computer screen.

A one-way ANOVA on comprehension across all four levels of typeface variation as presented as 8-bit text on a computer screen revealed no significant difference between scores of the two sample groups ($F = 1.46$, $df = 3, 84$) at the 0.05 level of significance. Therefore, hypothesis 4 could not be rejected.

Table 21

ANOVA Of Reading Comprehension Scores Across Four Levels Of Typeface Variation As They Were Displayed As 8-bit Text On A Computer Screen

| Source | <i>df</i> | <i>SS</i> | <i>MS</i> | <i>F</i> | <i>p</i> |
|--------|-----------|-----------|-----------|----------|----------|
| Form | 3 | 181.2 | 60.4 | 1.46 | 0.230 |
| Error | 84 | 3466.4 | 41.3 | | |
| Total | 87 | 3647.6 | | | |

=.05

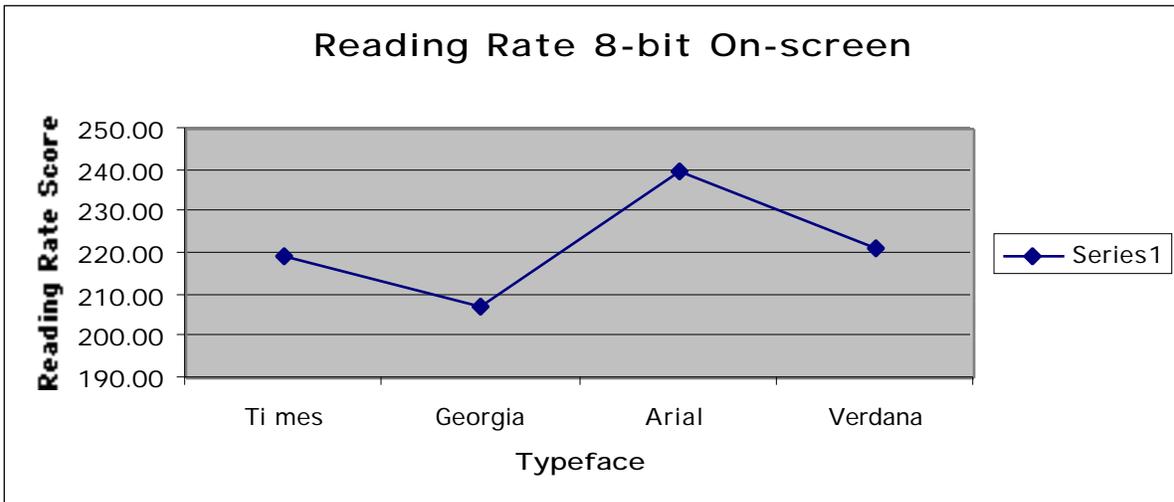


Figure 8. The results of the reading rate test across the four different levels of typeface as they were presented using 8-bit text on a computer screen.

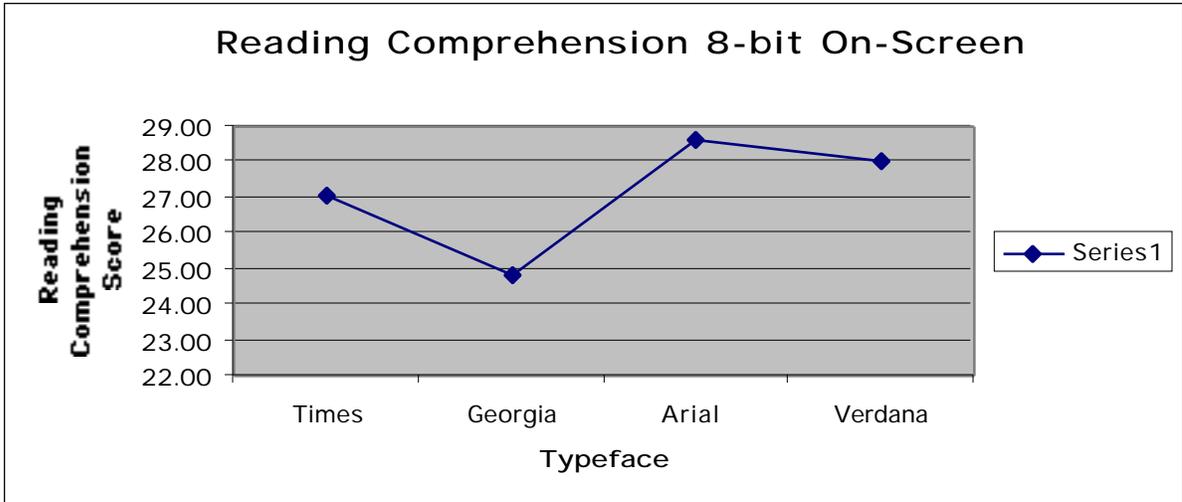


Figure 9. The results of the reading comprehension test across the four different levels of typeface as they were presented using 8-bit text on a computer screen.

Table 22

Descriptive Statistics of Reading Comprehensions Scores across Four Levels of Typeface Variation as They Were Displayed as 8-bit Text on a Computer Screen

| Level | <i>n</i> | <i>X</i> | <i>s</i> |
|---------|----------|----------|----------|
| Times | 22 | 27.091 | 5.554 |
| Georgia | 22 | 24.818 | 7.817 |
| Arial | 22 | 28.591 | 5.509 |
| Verdana | 22 | 28.000 | 6.539 |

Hypothesis 5: There was no significant difference between the mean scores of reading rate across four levels of typeface variation (Times, Georgia, Arial, and Verdana) as they were displayed as 600 dpi text on paper.

A one-way ANOVA on rate across all four levels of typeface variation as presented as 600 dpi text on paper revealed no significant difference between scores of the two sample groups ($F = 0.40$, $df = 3, 84$) at the 0.05 level of significance. Therefore, hypothesis 5 could not be rejected.

Table 23

*One-Way ANOVA of Reading Rate Scores across Four Levels of Typeface
Variation as They Were Displayed as 600dpi Text on Paper*

| Source | <i>df</i> | SS | MS | <i>F</i> | <i>p</i> |
|--------|-----------|--------|------|----------|----------|
| Form | 3 | 3612 | 1204 | 0.40 | 0.755 |
| Error | 84 | 254078 | 3025 | | |
| Total | 87 | 257690 | | | |

=.05

Table 24

*Descriptive Statistics of Reading Rate Scores across Four Levels of Typeface
Variation as They Were Displayed as 600dpi Text on Paper*

| Level | n | X | SD |
|---------|----|--------|-------|
| Times | 22 | 206.18 | 47.28 |
| Georgia | 22 | 201.86 | 75.23 |
| Arial | 22 | 190.45 | 53.90 |
| Verdana | 22 | 206.05 | 36.03 |

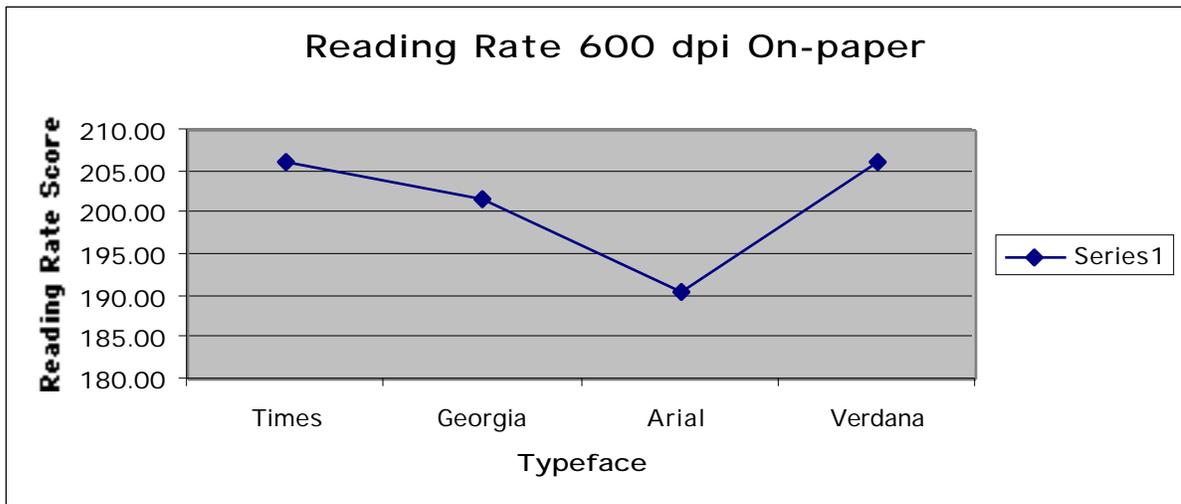


Figure 10. The results of the reading rate test across the four different levels of typeface as they were presented on 600 dpi paper output.

Hypothesis 6: There was no significant difference between the mean scores of reading comprehension across four levels of typeface variation (Times, Georgia, Arial, and Verdana) as they were displayed as 600 dpi text on paper.

A one-way ANOVA on comprehension across all four levels of typeface variation as presented as 600 dpi text on paper revealed no significant difference between scores of the two sample groups ($F = 0.40$, $df = 3, 84$) at the 0.05 level of significance. Therefore, hypothesis 6 could not be rejected.

Table 25

ANOVA of Reading Comprehension Scores across Four Levels of Typeface Variation as They Were Displayed as 600dpi Text on Paper

| Source | <i>df</i> | SS | MS | <i>F</i> | <i>p</i> |
|--------|-----------|--------|------|----------|----------|
| Form | 3 | 75.5 | 25.2 | 0.57 | 0.639 |
| Error | 84 | 3738.4 | 44.5 | | |
| Total | 87 | 3813.9 | | | |

=.05

Table 26

Descriptive Statistics for Reading Comprehension Scores across Four Levels of Typeface Variation as They Were Displayed as 600dpi Text on Paper

| Level | <i>n</i> | M | SD |
|---------|----------|--------|-------|
| Times | 22 | 27.273 | 5.513 |
| Georgia | 22 | 27.000 | 7.171 |
| Arial | 22 | 25.136 | 7.344 |
| Verdana | 22 | 27.455 | 6.501 |

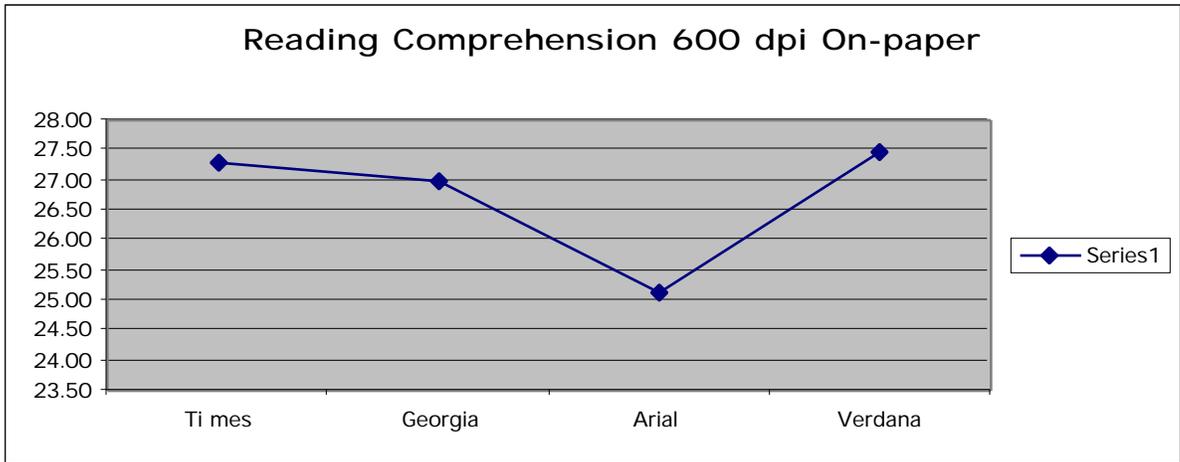


Figure 11. The results of the reading comprehension test across the four different levels of typeface as they were presented on 600 dpi paper output.

Other Analyses

Table 27

One-Way ANOVA Comparing Reading Rate Scores across Three Presentation Modes

| Source | <i>df</i> | <i>SS</i> | <i>MS</i> | <i>F</i> | <i>p</i> |
|--------|-----------|-----------|-----------|----------|----------|
| Mode | 2 | 31421 | 15710 | 4.16 | 0.017 |
| Error | 261 | 986740 | 3781 | | |
| Total | 263 | 1018161 | | | |

=.05

Table 28

Descriptive Statistics for Reading Rate Scores across Three Presentation

Modes

| Level | <i>n</i> | <i>X</i> | <i>s</i> |
|-----------------|----------|----------|----------|
| 1-bit On-screen | 88 | 197.00 | 61.86 |
| 8-bit On-screen | 88 | 221.93 | 67.47 |
| 600dpi On-paper | 88 | 201.14 | 54.42 |

Table 29

Tukey's Pairwise Comparisons for Reading Rate Scores across Three

Presentation Modes

| | 1-bit On-screen | 8-bit On-screen |
|-----------------|-----------------|-----------------|
| 8-bit On-screen | -46.6* | |
| | -3.2 | |
| 600dpi On-paper | -25.8* | -0.9 |
| | 42.5 | |

* $p < .05$ [$F_{.05}(2, 261) = 3.31$]

Table 29 shows a Tukey's pairwise comparison table indicating the individual differences between the groups. It shows that there is significant difference in reading rate between 1-bit text on screen and 8-bit text on screen, and, 1-bit text on screen and 600dpi text on paper, but found no significant difference between (2,3) 8-bit on-screen text and 600dpi text on paper.

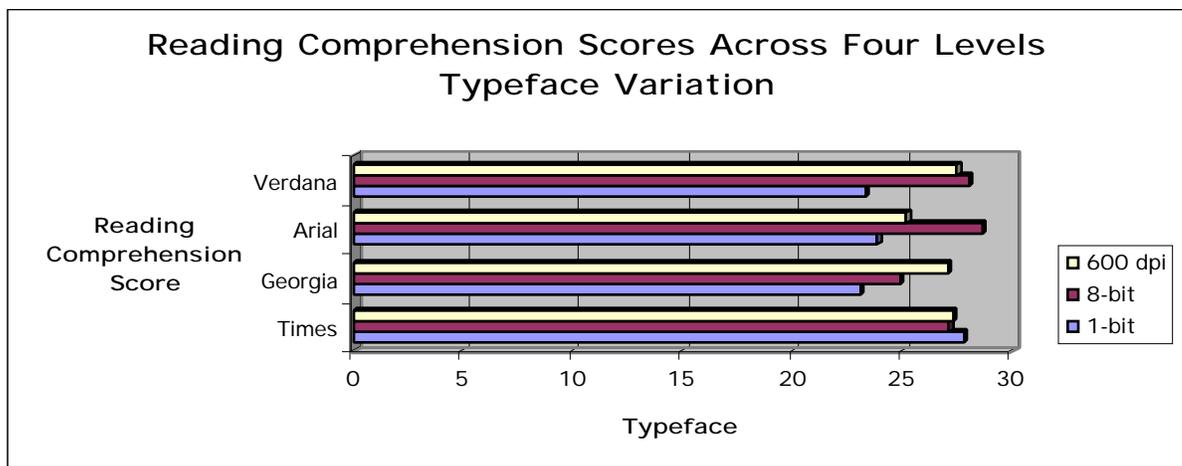


Figure 12. The results of the reading comprehension test across the different levels of typeface with respect to the mode of presentation.

Table 30

One-Way ANOVA Comparing Reading Rate Scores across the Three Presentation Modes in Times

| Source | <i>df</i> | SS | MS | <i>F</i> | <i>p</i> |
|--------|-----------|--------|------|----------|----------|
| Mode | 2 | 4959 | 2479 | 0.63 | 0.538 |
| Error | 63 | 249302 | 3957 | | |
| Total | 65 | 254260 | | | |

=.05

Table 31

Descriptive Statistics for Reading Rate Scores across the Three Presentation Modes in Times

| Level | <i>n</i> | <i>X</i> | <i>s</i> |
|-----------------|----------|----------|--------------|
| 1-bit on-screen | 22 | 198.09 | 49.26 |
| 8-bit on-screen | | 22 | 219.14 84.91 |
| 600dpi on-paper | 22 | 206.18 | 47.28 |

Table 32

Tukey's Pairwise Comparisons for Reading Rate Scores across the Three Presentation Modes in Times

| | 1-bit On-screen | 8-bit On-screen |
|-----------------|-----------------|-----------------|
| 8-bit On-screen | -66.5* | |
| | 24.4 | |
| 600dpi On-paper | -53.6* | -32.5* |
| | 58.4 | |

* $p < .05$ [$F_{.05}(2, 261) = 3.31$]

Table 32 shows a Tukey's pairwise comparison table showing the individual differences between the groups who read text set in Times across three presentation modes. It shows that there is significant difference between (1,2) 1-bit and 8-bit and (1,3), 1-bit and and between (2,3) 8-bit and 600dpi.

Above is a Tukey's pairwise comparison table showing the individual differences in comprehension score between the groups who read text set in Times across three presentation modes. It shows that there is significant difference in reading comprehension between (1,2) 1-bit and 8-bit and (1,3), 1-bit and between (2,3) 8-bit on-screen text and 600dpi text on paper.

Analysis Summary

Each of six previously stated hypotheses were tested during the data analysis portion of this study in an attempt to refute the null hypotheses. The data collected from twelve treatment groups were analyzed using Minitab statistical software which enabled a series of ANOVA operations as well as descriptive statistics to be quickly calculated. Initially, two, two-way ANOVA were run on the entire data set to determine if there existed any significant differences among the sample groups for both dependent variables—reading rate and reading comprehension across twelve levels of font and presentation mode variation. After determining the *F*-values for each treatment group and comparing them with their corresponding critical values, the only significance indicated in these tests was between the presentation modes used in the experiment. None of the six null hypotheses could be rejected.

Additional analyses using Tukey's test of pairwise comparison indicated significant differences between individual treatment groups. Consequently, significant difference was indicated between 1-bit on-screen and 8-bit on-screen in all cases, but—in some cases—not between 8-bit on-screen and 600dpi on-paper. Overall, less significance was found between treatment groups tested for the 8-bit on-screen and 600dpi on- paper presentation modes.

CHAPTER FIVE

Summary, Conclusions, and Recommendations

Background

One common way that people communicate and gather information today is through text-based means displayed on digital media such as CRT, LCD, and TV screens. This is a result of what many refer to as a “revolution” in communication technology that is currently facilitating the shift in textual information display from print media to computer screen. Widespread changes such as these call into question variables related to the readability of on-screen text. This study examined the use effect of fonts designed specifically for computer displays on reading rate and reading comprehension.

Typefaces designed for the screen are a relatively new innovation. They represent an attempt to compensate for the lack of resolution of the bitmaps of CRTs and for the fact that the majority of fonts available today were designed to be read on paper. The use of these older fonts on computer displays forces one to read “type designed in another era for another medium” (Will-Harris, 1998).

In the past, readability research has examined printed text on paper, and typography designed for print as it has been rendered on screen. In the majority of studies it has been shown that reading performance of subjects who read

text presented on paper is much better than that of subjects reading from alternative presentation media. However, relatively little research has been done to examine the readability of typography designed to be read directly on computer screens.

Previous research in the area of readability has focused upon the effects of typographic variables such as typeface (e.g. serif versus sans serif typefaces), letter spacing, line spacing (or leading), justification contrast, resolution, inverted text, mechanically-tinted backgrounds, type size, type style, and letter spacing. Other research has examined the media or display conditions upon which text is displayed as well as demographic factors of the readers themselves on reading rate and comprehension. Although research exists which examines the differences between the readability of textual material on-paper and on-screen, the rapid evolution of digital typographic tools and display technologies bear periodic/updated examination due to the risk of setting inefficient standards.

The literature surrounding the study of readability indicates a lack of research that substantiates the advantages of using screen-optimized fonts for the express purposes of textual display on a computer screen. These commercially available fonts are touted by Microsoft as "more readable" for the purposes of screen display of text. Most commonly they are employed in the creation of web pages, e-books, or e-mail messages. However, some research has indicated their appropriateness for use in the on-screen display

of Adobe Portable Document Files (PDFs), another commonly used mode of networked document delivery (Mather, 1997).

Purpose

The purpose of this study was to determine empirically whether the text displayed using the typefaces Georgia and Verdana [which, according to Microsoft, 1998; Will-Harris, 1998; Monotype, 1997, optimize on-screen readability] significantly improve reading performance as measured by reading rate and/or reading comprehension as compared to the typefaces Times and Arial—both of which are digital fonts designed specifically for text-on-paper output.

Procedure

In order to empirically compare the reading performance of subjects who read text set in two different typefaces (Georgia and Verdana) which were specifically designed for use on computer screens as compared with two typefaces which were specifically designed for print output (Times and Arial), an experiment was conducted in which subjects were tested for reading rate and reading comprehension using a standard, validated test instrument (the *Nelson-Denny Reading Test*). The tests were administered using test forms presented either in one of three different presentation modes on either a computer screen or on paper. Also, subjects were given versions of the test set in one of the typefaces under examination (Georgia, Verdana, Times, or Arial)

and were asked to take the test as it was displayed in one of three possible presentation modes: 1-bit text on-screen, 8-bit text on-screen, or 600dpi laser-printed output. As a result, the entire sample selected for this study was divided into twelve distinct treatment groups—each being tested for a different combination of typeface and/or media type. Those group combinations are listed in *Table 33*.

Table 33

Treatment Groups

| | |
|-----------------|---|
| Group 1 | read 1-bit text set in the font Georgia on a computer display |
| Group 2 | read text rendered at 600dpi set in the font Georgia on paper |
| Group 3 | read 8-bit text set in the font Georgia on a computer display |
| Group 4 | read 1-bit text set in the font Times on a computer display |
| Group 5 | read text rendered at 600dpi set in the font Times on paper |
| Group 6 | read 8-bit text set in the font Times on a computer display |
| Group 7 | read 1-bit text set in the font Verdana on a computer display |
| Group 8 | read text rendered at 600dpi set in the font Verdana on paper |
| Group 9 | read 8-bit text set in the font Verdana on a computer display |
| Group 10 | read 1-bit text set in the font Arial on a computer display |
| Group 11 | read text rendered at 600dpi set in the font Arial on paper |
| Group 12 | read 8-bit text set in the font Arial on a computer display |

Findings

Six hypotheses were tested. The data collected from the twelve treatment groups (See Table 29) were analyzed using Minitab™ statistical software.

Initially, two, two-way analyses of variance were run on the entire data set to determine if there existed any significant differences among the sample groups for both dependent variables—reading rate and reading comprehension across twelve levels of font and presentation mode variation. After determining the F -values for each treatment group and comparing them with their corresponding critical values, the only significance indicated in these tests was between the presentation modes used in the experiment. None of the six null hypotheses could be rejected.

Additional analyses were conducted to identify between-group differences. Using Tukey's test of pairwise comparison, significant differences between individual treatment groups could be differentiated. Consequently, significant difference was indicated between 1-bit on-screen and 8-bit on-screen in all cases, but—in some cases—not between 8-bit on-screen and 600dpi on-paper. Overall, less significance was found between treatment groups tested for the 8-bit On-screen and 600dpi presentation modes. Major findings of the data analysis are listed below:

1. There was no significant difference between the reading rates of sample groups who read text from the four different levels of variation in font ($F = 0.03$, $df = 3, 252$) at the 0.05 level of significance.
2. There was no significant interaction between font and presentation mode in respect to reading rate ($F = 1.07$, $df = 6, 252$) at the 0.05 level of significance.

3. Significant difference was found among the reading rates of sample groups who read text from the three different levels of variation in presentation mode ($F = 4.12$, $df = 2$, 252) at the 0.05 level of significance.
4. There was no significant difference between the reading comprehension scores of sample groups who read text from the four different levels of variation in font ($F = 0.62$, $df = 3$, 252) at the 0.05 level of significance.
5. There was no significant interaction among font and presentation mode with respect to reading comprehension ($F = 0.81$, $df = 6$, 252) at the 0.05 level of significance.
6. Statistical significance was found between the reading comprehension scores of sample groups who read text from the three different levels of variation in presentation mode: 1-bit on-screen, 8-bit on-screen, and 600dpi on-paper ($F = 6.97$, $df = 2$, 252) at the 0.05 level of significance. See *Table 7* in Chapter Four for a summary table of the ANOVA for reading comprehension.
7. Statistical significance was found between the reading rates of sample groups who read text from the three different levels of variation in presentation mode: 1-bit on-screen, 8-bit on-screen, and 600dpi on-paper ($F = 4.12$, $df = 2$, 252) at the 0.05 level of significance.
8. There was no significant difference in reading rate scores between 8-bit on-screen text and 600dpi text on paper ($F = .9$, $df = 2$, 261) at the 0.05 level of significance.

Conclusions

The conclusions derived from the this study are as follows:

1. Since it was found that 8-bit on-screen text was not significantly more readable than 600dpi text on paper, and 1-bit on-screen text was found to be significantly *less* readable 8-bit on-screen text and 600dpi text on paper, this research concludes that for purposes of ease of readability, on-screen text is better suited to be rendered as 8-bit on-screen text than 1-bit on-screen text. Also, the findings indicate that 8-bit on-screen text was not found to be significantly less readable than 600dpi text on paper.
2. Due to the various typefaces currently being used in digital typography and the differing presentation media, further exploration of the readability of on-screen text should examine more fonts and screen display variables.

Observations

1. The font Georgia does not significantly increase reading speed or reading comprehension as compared with its designed-for-print, serif counterpart Times when displayed on a computer screen as either as 1-bit on-screen text, or 8-bit on-screen text.
2. The font Verdana does not significantly increase reading speed or reading comprehension as compared with its designed-for-print, sans serif

counterpart Arial when displayed on a computer screen either as 1-bit text on-screen, or 8-bit on-screen text.

3. Reading rate scores of group who read text set in Times were higher—but not significantly higher—for the groups when read from 8-bit on-screen text than the group who read from 660dpi on-paper text.
4. Reading rate scores for those who read 8-bit on-screen text set in Times was higher—but not significantly higher—than for those who read 600dpi on-paper text from all other typefaces examined in this study (Georgia, Arial, and Verdana).
5. Reading rate scores for those who read 1-bit on-screen text set in Georgia were higher—but not significantly higher—than for those who read 1-bit on-screen text set in all other typefaces examined in this study (Times, Arial, and Verdana).
6. Reading rate scores for those who read 8-bit on-screen text set in Arial were higher—but not significantly higher—than for those who read 8-bit on-screen text set in all other typefaces examined in this study (Times, Georgia, and Verdana).
7. Reading comprehension scores for those who read 600dpi on-paper text set in Verdana were higher—but not significantly higher—than for those who read 600dpi on-paper text set in all other typefaces examined in this study (Times, Georgia, and Arial).

8. Reading comprehension scores for those who read 8-bit on-screen text set in Arial were higher—but not significantly higher—than for those who read 8-bit on-screen text set in all other typefaces examined in this study (Times, Georgia, and Verdana).
9. Reading comprehension scores of groups who read 1-bit on-screen text set in Times were higher—but not significantly higher—than for those who read 1-bit on-screen text set in all other typefaces examined in this study (Georgia, Arial, and Verdana).

Discussion

A review of the literature surrounding the Microsoft fonts Georgia and Verdana suggests that the implementation of these fonts when typesetting documents expressly for the computer screen enhances the readability of those documents (Microsoft, 1998; Will-Harris, 1998; Monotype, 1997). The creators of these fonts, Matthew Carter and Tom Rickner, developed these fonts from the "ground-up" for the computer screen taking into primary consideration the bitmap composition which is characteristic of computer displayed text. These fonts were created to fulfill the need that Microsoft Corporation expressed for fonts which facilitate readability for web pages and documents displayed on computer screens. In the words of Microsoft's Director of Typography Bill Hill, "As the need for viewing on-screen documents increases, Microsoft desired to produce a set of highly legible fonts that our customers could use specifically for this purpose" (Monotype, 1997). Embodied

in the Georgia and Verdana are fonts which are designed to enhance the readability of on-screen documents as well as printed documents (Microsoft, 1998).

The four typefaces tested in this study were chosen because of their standard usage in word processing and electronic publishing. Microsoft virtually sets the standard for word processing due to their size and overwhelming influence in the business and productivity software areas. This dominance practically bestows Microsoft with the power to impose standards for typography by building its own specifications into its software products. Guarding against technological mandates from commercial giants such as Microsoft (which could potentially be flawed at worst, or, at least, not optimal) is in the interest of the population of readers of on-screen documents.

Although careful to control for all extraneous variables, some environmental conditions of this test related to real-life situations as opposed to more clinically-controlled environments of other experimental studies. These testing conditions may have had some influence on the results of the study. However, in the interest of applying real-life situations, environmental surroundings were uniform throughout the study. A common phenomena noted by Kennedy and Murray (1993, p.258), involves the possibility that some of the subjects may have developed skill in performing specialized tasks to which they may have grown more accustomed to than other subjects. Respondents were asked to report an approximate amount of time (per week) that they

typically spent reading from computer displays. Responses to this query varied widely.

Passage difficulty should not have been a factor affecting any of the subjects used in this study. The documentation for the *Nelson-Denny Reading Test* stated that the passages used for Form G were validated for students from 9th to the 16th level of reading ability. In addition the latest revisions of the *Nelson-Denny Reading Test* reflected in the *Forms G and H*, made an attempt to account for racial biases by adapting the content of the textual passages to reflect a more diverse population. Student attitudes toward the passages may have come into play. Although the duration of the testing situation was less than thirty minutes, some subjects displayed apathetic attitudes toward the test itself. This could have potentially negatively affected their scores on the test.

The observations extracted from the data gathered in this study lead one to believe that the screen-optimized typefaces (Georgia and Verdana) do not necessarily improve readability. These observations point out where some obvious differences in reading performance occurred. These differences, however, do not reflect any statistically significant difference. Therefore, they should be taken as purely "observational".

Surprisingly, reading rate scores of the group who read text set in Times were higher for the groups when read from 8-bit computer display text than the group who read from paper. As displayed as 1-bit text on a computer screen, reading rate did increase. However it did not differ significantly. Subjects reading

Georgia averaged 213.59 words per minute (wpm), whereas those reading Times averaged 198.09 wpm. Curiously, when both displayed on screen as 1-bit text, Times outperformed Georgia although not significantly. Over a one-minute time span, those reading text in Times averaged 219.14 wpm, whereas those reading text in Georgia read only 207.55 wpm. On paper, Times outperformed Georgia 206.18 wpm to 201.86 wpm.

For reading comprehension, both Times and Arial outperformed Georgia and Verdana for text displayed at 1-bit on screen although the differences were not significant. This runs counter to the assertion by Microsoft that their fonts are more readable on-screen.

For reading rate and reading comprehension, both Times and Arial outperformed Georgia and Verdana at 8-bit on-screen. This also runs counter to the assertion by Microsoft that their fonts are more readable on-screen.

When displayed as 8-bit text on a computer screen, both reading rate and reading comprehension scores are highest when text is set at font Arial, 12-point as compared with fonts Times, 12-point, Georgia, 12-point, and Verdana, 12-point. Reading speed was observed to be significantly greater in those treatment groups who read text from paper as opposed to those who read text from a computer screen. Reading comprehension was observed to be significantly greater in those treatment groups who read text from paper as opposed to those who read text from a computer screen.

According to the above observations, this research cannot statistically substantiate claims made by Microsoft that advocate the use of their WebFonts for the purposes of increased readability. Although some observations lead one to believe that these fonts (Georgia and Verdana) positively affect reading rate and reading comprehension as compared with other fonts (Times and Verdana), the findings of this study indicate no significant advantage of the use of these fonts in any of the situations examined herein.

The foremost conclusions drawn from this experiment are the indication that reading text displayed as 8-bit text on a computer screen does not significantly slow reading speed as compared with 600dpi laser printed paper output. Evidence of this nature has not generally been found in prior research comparing on-screen reading with on-paper reading. According to the data in this study, this result is most likely a consequence of higher quality of the rendering of the on-screen text rather than the typographic design of typefaces designed for the computer screen.

Recommendations

1. Limited generalizations have been noted as a weakness in typographic research (Holmes, 1986). The rapidity of technological advance in text-based digital media will certainly not aid in this shortcoming. Periodic research examining the readability of typographic innovations would lend direction to further development in this area. Since this research examined

a sample of students enrolled in specialized curriculum (the College of Applied Science and Technology at Illinois State University), the wider generalizations could be obtained by repeating the experiment in other settings and/or concentrations of study (e.g.: students enrolled in vocational/technical schools, liberal arts colleges and universities, art institutes, etc.)

2. Increased consumer and educator awareness of text readability would naturally improve the efficacy of text-based design and information display. This is especially true in an age when worldwide publishing is easily accessible from the desktop.
3. Unbiased typographic testing standards should be encouraged. Non-biased testing would allow for a more honest approach to designing and marketing technological tools.
4. No specialized instrument currently exists for testing the readability of typefaces. The test used in this research was one of many tests used in readability research. Therefore, a specialized standard test for measuring typographic variables should be developed and standardized.
5. Many studies of readability take into consideration the subjective preferences of the reader (Canary, 1983; Turner, 1982; Tullis, Boynton, & Hersh's, 1995; Taylor, J. L., 1990; O'Regan, K., N. Bismuth, Hersh, R.D., & Pappas, A., 1998). Although not a variable of ergonomic research, this variable would be appropriate for study in a behavioral study of readability

due to the vastness of the customization capabilities available in computing environments offered by existing and potential machines.

6. Due to the implications of this research, it is suggested that those responsible for publishing textual documents online should strongly consider using 8-bit text for display or distribution of those documents (e.g. electronic books, electronic theses and dissertations, etc.). As indicated by the findings of this research, the font Arial 12-point is recommended for optimization of reading speed and reading comprehension.
7. Time of day may have been a contributing factor to the performance of the subjects. This test was administered to groups of subjects at various Times between 8 a.m. and 7 p.m., however, time of day was not taken into consideration as an extraneous variable in this study. Therefore, it is recommended that in future studies the variable of time of day be held constant, or, at least, taken into consideration in the analysis of the reading performance measures of speed and comprehension.
8. Since the reading test administered in this research was timed (twenty-minute total), the dependent variable of reading comprehension was in some cases affected by a subjects reading rate. Due to the fact that—in some cases—lack reading rate prohibited subjects from finishing the test of comprehension resulting in incomplete (lower) comprehensions scores. For future research, it is recommended that the test be administered without the twenty-minute time limit.

Implications for Practice

1. The combined results of both the reading rate and reading comprehension portions of the test suggest 12-point Arial appears to be better suited for text displayed at 8-bit on-screen if this text is meant to be read directly from the screen (in Adobe's PDF format, for example).
2. The combined results of both the reading rate and reading comprehension portions of the test suggest 12-point Georgia appears to be better suited for text displayed at 1-bit on-screen if this text is meant to be read directly from the screen.
3. The combined results of both the reading rate and reading comprehension portions of the test suggest 12-point Verdana or 12-point Times appears to be better suited for text displayed on 600dpi laser-printed output.

REFERENCES

Acrobat Systems Incorporated (1998). Adobe web type. Retrieved June 6, 1998 from the World Wide Web: <http://www.adobe.com/prodindex/webtype/main.html>

Anstis S. M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research*, 14, 589-582.

Babbie, E. (1990). *Survey research methods* (2nd ed.). Belmont, California: Wadsworth.

Berryman, G. (1984). *Notes on graphic design and visual communication*. Los Altos, California: Kaufman.

Beach, M. (1992). *Graphically speaking: an illustrated guide to the working language of design and printing*. Manzanita, Oregon: Elk Ridge.

Bond, M. (1998). Suite101.com. Retrieved from the WWW on Tuesday, June 16, 1999 from <http://www.suite101.com/articles/article.cfm/1329>

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226, 177-178.

Brown, J. I., Fishco, V. V., & Hanna G. (1993). *The Nelson-Denny Reading Test*. Itasca, IL: Riverside Publishing.

Buckingham, B. R. (1931). *New data on the typography of textbooks*. Thirteenth yearbook of the national society for the study of education. Bloomington, IL: Public School Publishing Company, 93-125.

Burger, J. (1993). *The desktop multimedia bible*. Reading, Massachusetts: Addison-Wesley.

Burke, S. (1998, January). Microsoft working on web readability. *Computer Reseller News*. Retrieved November 17, 1998 from the World Wide Web: <http://www.crn.com/dailies/weekending112098/nov17dig10.asp>

Canary, M.R. (1983). *An analysis of typographic factors on the visual information processing of disabled readers*. Unpublished doctoral dissertation, The American University, Washington, D.C.

Cartwright, D., Jr. (1991). *The effects of desktop publishing on typography from the perspective of professional typographers*. Unpublished masters thesis, East Texas State University, Commerce.

Cavanaugh, S. (1995). *Digital type design guide*. Indianapolis: Hayden.

DiNucci, D. (1995, January). Future fonts. *Print*, 49, 28-36.

Gillard, A. R. (1923). The effect on reading of changes in the size of type. *The Elementary School Journal*, 24, 138-146.

Gerig, J., Nibbelink, W. H., & Hoover, H. D. (1992, December). The effect of print size on reading comprehension. *Iowa Reading Journal*, 3, 26-28.

Glisky, E.L. (1984). *Encoding and retrieval effects in memory for inverted words*. Unpublished doctoral dissertation, University of Toronto.

Gould, J. (1981). Composing letters with computer-based text editors. *Human Factors*, 23, 593-606.

Gould, J., & Grischkowsky, N. (1984). Doing the same work with hard copy and with cathode-ray tube (CRT) computer terminals. *Human Factors*, 26, 323-337.

Graphion Typesetting & Systems, Inc. (1997). Elements of typesetting style. Retrieved November 11, 1997 from the World Wide Web:
<http://www.slip.net/~graphion/legnote.html>

Gulick, R. (1996). Vendors vie for web font standard. *MacWeek* 10, 12-13.

Haas, C., & Hayes, J. R. (1985). Effects of text display variables on reading tasks: computer screen vs. hard copy. *Communications Design Centre Technical Reports Series*. Pittsburgh: Caregie-Melon University.

High, C. (1998). The art of design for legibility. Retrieved February 2, 1998 from the World Wide Web: <http://www.htalcyon.com/chigh/artlegw.html>

Hobbes, R. (1998). Hobbes' Internet timeline. Retrieved October 15, 1998 from the World Wide Web:
<http://info.isoc.org/guest/zakon/Internet/History/HIT.html>

Holmes, J. A. (1986). *Formatting variables and typeface variations of dot-matrix print and their effect on reading comprehension and reading rate*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg.

Howell, D.C. (1997). *Statistical methods for psychology* (4th ed.). Belmont, California: Duxbury Press.

Illinois State University (1998). Illinois State University undergraduate catalog: 1998-99. Illinois State University, Normal, Illinois.

Johnson, K. (1998). Readability and reading ages of school science books. Retrieved March 30, 1998 from the World Wide Web:
<http://www.timetabler.com/reading.html>

Jonassen, D. H. (1982). *The technology of text: principles for structuring, designing, and displaying text*. Englewood Cliffs, NJ: Prentice-Hall.

Jonassen, D. H. (1985). *The technology of text volume two: principles for structuring, designing, and displaying text*. Englewood Cliffs, NJ: Prentice-Hall.

Journa, G. (1989). *Effects of image quality on reading performance and perceived image quality from CRT and hard-copy displays*. Unpublished masters thesis, Virginia Polytechnic Institute and State University, Blacksburg.

Kennedy, A. & Murray, W. S. (1993). Display properties and eye-movement control. In *studies in visual information processing, perception and cognition: advances in eye movement research*. d'Ydewalle, G. & Van Rensbergen, J. (Eds.) Amsterdam: North-Holland.

Kleper, M.L. (1987). *The illustrated handbook of desktop publishing and typesetting*. Blue Ridge Summit, Pennsylvania: TAB Professional and Reference Books.

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

Lancaster, F.W., & Warner A. (1985). Electronic publication and its impact on the presentation of information. In Jonassen, D. H.. *The technology of text volume two: principles for structuring, designing, and displaying text*. Englewood Cliffs, NJ: Prentice-Hall.

Le Rohellec, J., Brettel, H., & Glassar, J. (1996). LCD legibility under different lighting conditions as a function of character size and contrast. *Society for Information Display '96 Digest*, 35, 821-824.

Levin, J. R. & Kaplan, S. A. (1972, October) Imaginal facilitation of paired-associate learning: a limited generalization. *Journal of Educational Psychology* 63, 429-432.

Lighthouse Research Institute (1998). Studies in reading and legibility. Retrieved March 30, 1998 from the World Wide Web:
<http://www.lighthouse.org/1lh22t.html>

Lough, J.L. (1982). *Readability: an experiment to test the impact of selected graphic dimensions*. Unpublished masters thesis, University of Georgia.

Mather, P. (1997). *From paper to screen: revision electronic theses and dissertation guidelines for on-screen viewing*. Unpublished manuscript, Virginia Polytechnic Institute and State University, Blacksburg.

Microsoft Corporation (1998). Typography on the web. Retrieved on February 23, 1998 from the World Wide Web:

<http://www.microsoft.com/typography/web/default.htm>

Monotype (April, 1997). Monotype produces web fonts for Microsoft: new fonts integral to Internet Explorer. Retrieved on June 19, 1998 from the World Wide Web: http://www.monotype.com/html/news/nr_ms_web.html

Netscape Corporation (1998). Netscape Navigator: Version 4.05. Mountain View, California: Netscape Communications Corporation.

Nazir, T. A., Jacobs, A. M., O'Reagan, J. K. (1998) Letter legibility and visual word recognition. *Memory & Cognition*. 26 (4), 810-821.

Nazir T. A. , Heller D., & Sussmann C. (1992). Letter visibility and word recognition: The optimal viewing position in printed words. *Perception & Psychophysics*. 52, 315-328.

Nazir T. A., O'Reagan J. K., & Jacobs A. M. (1991). On words and their letters. *Bulletin of the Psychonomic Society*, 29, 171-174.

Olzak, L.A. & Thomas J.P. (1986). Seeing spational patterns. In K.R. Boff, Kaufman, & J. P. Thomas (Eds.) , *Handbook of perception and human perfomance*. New Youk: Wiley.

O'Regan, K., N. Bismuth, Hersh, R.D., & Pappas, A. (1998). *Legibility of perceptually-tuned grayscale fonts*. Paris: Laboratoire de Psychologie Expérimentale & Lausanne: École Polytechnique Fédérale.

Page, B. & Holm, D., (1996). *Web publishing with Adobe Acrobat and PDF*. New York: Wiley Computer Publishing.

Pedhazur, E.J. & L.P. Schmelkin (1991). *Measurement, design, and analysis: an integrated approach*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.

Pfaffenberger, B. (1996). *Publish it on the web*. Boston: Academic Press.

PR Newswire (1998). American Internet user survey finds more than 41.5 million U.S. adults are actively using the Internet. Retrieved on July 3, 1998 from the World Wide Web:

<http://www.prnewswire.com:80/cgi-bin/stories.pl?ACCT=104&STORY=/www/story/1-27-98/402429&EDATE=>

Prust, Z.A. (1989). *Graphic Communications: the printed image*. South Holland, Illinois: Goodheart-Wilcox Company, Inc.

Riverside Publishing (1998). *The Nelson Denny Reading Test*. Retrieved on December 18, 1998 from <http://www.riverpub.com/>

Rogers, D. J. (1986). *A study of the use of the American College Testing Program, Iowa Silent Reading Test, And Nelson-Denny Reading Test as predictors of the grade point averages of special admissions freshmen*. Unpublished doctoral dissertation, Southern Illinois University, Carbondale.

Taylor, J. L. (1990). *The effect of typeface on reading rates and the typeface preferences of individual readers*. Unpublished doctoral dissertation, Wayne State University.

Townsend, J. T., Taylor, S. G., & Brown D .R. (1971). Lateral masking for letters with unlimited viewing time. *Perception & Psychophysics*, 10, 375-378.

Turner, R. J. (1982). *An experimental study of the legibility of minus letterspacing of typography as measured by rate of reading and comprehension*. Unpublished doctoral dissertation, Ohio University.

Tullis, T.S., J. L. Boynton, & H. Hersh (1995). Readability of fonts in the Windows environment. CHI '95 Proceedings. Retrieved March 2, 1998 from the World Wide Web: http://www.acm.org/sigchi/chi95/proceedings/intpost/tst_bdy.htm

Uhl, W. L. (1937). *The materials of reading. thirty-sixth yearbook of the national society for the study of education*. Bloomington, IL: Public School Publishing company.

W3C (1997). Web style sheets. Retrieved on February 16, 1997 from the World Wide Web: [:http://www.w3.org/Style/](http://www.w3.org/Style/).

Walraven, J. (1992). Color basics for the display designer. In Widdel, H., & D. Post (Eds.), *Color in electronic displays*. New York: Plenu Press.

Weinman, L. (August, 1997). Visual designer: type on the web. *Web Techniques*, 2, 8 - 9.

Will-Harris, D. (1998). Georgia & Verdana: Typefaces designed for the screen (finally). *Typofile Magazine*. Retrieved on June 16, 1998 from the World Wide Web: <http://www.will-hariss.com/Verdana-Georgia.htm>

Will-Harris, D. (1998). The best faces for the screen. *Typofile Magazine*.
Retrieved on June 16, 1998 from the World Wide Web: <http://www.will-hariss.com/typoscrn.htm>

Wright, P., & Lickorish, A. (1983). Proof-reading on screen and paper.
Behaviour and Information Technology, 2, 227-235.

APPENDIX A:

Typography Primer

A typeface is a character set which shares a similar appearance and design. The character set includes letters, numbers, and symbols. Traditionally, the term “font” represents a complete set of characters or symbols, which share the same size and style. For example: 10 pt. Palatino Bold, 12 pt. Palatino Bold, and 10 pt. Palatino Ital. are all different fonts, but the same typeface is used in each.

There exist six generally recognized classifications by which all typefaces can be categorized. These typestyle categories include Roman typefaces, Sans Serif typefaces, Text typefaces, Script typefaces, Occasional typefaces, and Square Serif typefaces.

1. **Roman typefaces** always include serifs and stroke variations. This is the most commonly used typeface for body text (text set at 12-points or less). Examples of Roman typefaces are Times, Georgia, and Palatino.
2. **Sans Serif** typefaces do not have serifs. They are generally more modern looking than traditional Roman typefaces. They also have a uniform stroke width as opposed to Roman typefaces variable stroke width. Examples of sans serif typefaces are Arial, Geneva, and Verdana.

3. **Square Serif** typefaces are sometimes referred to as “Egyptian” typefaces.

They are most commonly set in large point sizes, due to the fact that copy set in this typeface appears very dense and confusing when set in smaller, body text point sizes. Square Serif typefaces are often used to communicate a sense of strength and power. Examples of square serif typefaces are Lubalin and City.

4. **Text** typefaces have the look of handwriting of medieval scribes. Most typographers agree that for legibility's sake these typefaces should never be set in all caps. Examples of text typefaces are London and Old English.

5. **Script** typefaces imitate the “cursive” style of handwriting. The individual characters of this typestyle category are often stylistic with joining characters. Examples of script typefaces are Nupital Script and Zapf Chancery.

6. **Occasional** typefaces generally include anything that cannot be otherwise classified in the other five typestyle categories. These typefaces are commonly known as novelty, or decorative typefaces. Examples of occasional typefaces are Dingbats, WingDings, Pi Characters.

Type Family

A type family includes all the various styles within a typeface. For example: Helvetica Bold, Helvetica Italic, Helvetica Bold Italic, Helvetica Condensed, Helvetica Expanded, etc.

Printers Measurement System

- A. 12-points = one pica
- B. 6 picas = one inch
- C. 72 points = one inch

Points

- A. Smallest increment of typographic measure
- B. Used to measure height of type, ruled line thickness

Picas

- A. 12-points equal one pica
- B. Used to measure line width

Measurement of Type

- A. Common Point sizes: 7, 8, 9, 10, 11, 12, 14, 18, 24, 36, 48, 60, 72, 96
- B. All text set 12 pt. or below is called body text.
- C. All text set at or above 14 pt. is called display text.
- D. Line Space (leading) is the space (measured in points) between rows of text, from baseline to baseline.
- E. When the line spacing equals the point size, type is said to be “set solid”.

F. Measurement of Type

G. Line Length is measured in picas.

Justification

A. left justified = quad left = right justified

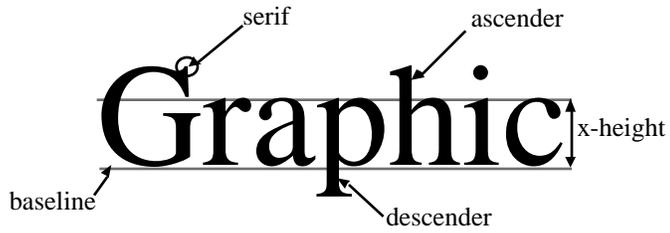
B. ragged right = quad right = ragged left

C. (Fully) Justified = Flush right and left

Kerning

Purposely reducing the space between certain letters to compensate for esthetically objectionable space.

The Anatomy of Text



1. x-height (body height)
2. type rests on the baseline
3. ascenders and descenders
4. stroke: variable and uniform

APPENDIX B:

Instructions for the Reading Test

Instructions for Reading Test:

Instruction: Distribute answer sheets and have students begin filling out the demographic data. You may take this test in pencil or in pen. Raise your hand if you do not have a pencil or pen to write with today."

Read aloud: Is there anyone in this room under the age of 18 years.

Read aloud: "Today you are going to take a test that will help us find out how well you read text presented on paper or on a computer screen. The data that is collected from this test will be analysed as a *group* to determine how reading performance is affected by the mode of presentation."

Read aloud: **DISCLAIMER:** "Please be aware that your participation in this test is strictly voluntary. If you do not wish to participate in this test you are free to leave now or at anytime during the test. There is no penalty for not participating or completing any items on this test."

Read aloud: "Right now I am going to assign you a number between one and twelve. Remember this number. The number that you are

assigned will tell you which test form you will be using. If your number is between one and eight, you will take the test from a computer screen. Please write the number of your test form in the box in the upper-left-hand corner of the test answer sheet. If you receive a paper test booklet please do not open it until I instruct you to do so. Also, do not write on the answer sheet"

Instruction: "If you have test forms 1, 2, 3, or 4, first open the file and maximize the document to fill the entire screen by pressing the "maximize" button in the upper-right-hand corner of the document window [Demonstrate] Please read the directions on the first page. Please pull down the "View" menu and select "Zoom". Next, type 90 into the "Percent" field. DO NOT PROCEED UNTIL INSTRUCTED TO DO SO.

Read aloud: Those of you with test forms 5, 6, 7, or 8, maximize the document to fill the entire screen by pressing the "maximize" button in the upper-right-hand corner of the document window [Demonstrate] Please read the directions on the first page. DO NOT PROCEED UNTIL INSTRUCTED TO DO SO.

Read aloud: I will now pass out the test booklets for those of you who were randomly chosen to receive test forms 9, 10, 11, or 12. Leave the booklets closed, and do not mark on the answer sheets until I tell you to do so. If there are lines surrounding the sample text, please pull down the "Table" menu and select "Hide Gridlines". DO NOT PROCEED UNTIL INSTRUCTED TO DO SO.

Instruction: "If you have test forms 5, 6, 7, or 8, first open the file. Please pull down the "View" menu and select "Zoom To...". Next, type "120" into the "Magnification" field. Read the directions on the first page. DO NOT PROCEED UNTIL INSTRUCTED TO DO SO.

Instruction:"If you have test forms 1, 2, 3, or 4, pull down the "VIEW" menu and deactivate all toolbars that are currently activated.

*Instruction:*When I tell you to begin, please use the "Pg Dn" and "Pg Up" buttons on the keyboard only to navigate the document." When I tell you to begin, please use the "Pg Dn" and "Pg Up" buttons on the keyboard only to navigate the document."

Read aloud: "Are there any questions?"

Read aloud: This test is a reading comprehension test containing 38 items.

Your score is based on the number of *correct* responses. Since there is no penalty for incorrect answers, it is to your advantage to mark every question you read. But do not spend too much time on any one question.

Read aloud: Please do not turn this page of the test booklet or scroll-down to other pages of the document until directed to do so.

Instruction: When I tell you to begin the test, those of you with paper booklets will open to the second page of the test and begin reading the text passage. At the same time, those of you whose test appears on the computer screen will use the Pg Dn key to navigate to the second page in the document and begin reading the text passage.

Instruction: During the entirety of this test, I would like you to navigate the document using **ONLY** the Pg Dn and/or Pg Up buttons located on the numerical keypad of your keyboard. [Demonstrate]. Do not use any other keys on the keyboard except the Pg Up and/or Pg Dn keys.

APPENDIX C

Courses in the Department of Industrial Technology
from which the subjects in this subject were sampled

IT100 - Introduction to Industrial Technology

IT211 - Architectural Drafting

IT 150 - Fundamentals of Printing Technology

IT120 - Introduction to Building Construction

IT121 - Construction Materials and Methods

IT250 - The Graphic Arts Processes

IT283 - Technology of Information and Imaging Systems

IT323 - Construction Estimating and Bidding

IT325 - Construction Scheduling and Finance

IT225 - Construction Equipment Management

ACS155.02 - Introduction to Microcomputers

IT 151 - Introduction to Industrial Computing Systems

IT383 - Telecommunications Technology

APPENDIX D:

Answer Sheet

Test Form Number:

In this box,
please write the number of the
test form from which you read.

Before you begin, please answer the following questions:

1. What is your major? _____
2. What is your gender? Male Female
3. Do you wear corrective lenses? Yes No
4. How old are you? _____
5. What is your academic classification?
 Freshman Sophomore Junior Senior Graduate
6. What is your racial/ethnic group:
 American Indian - Alaska Native Asian or Pacific Islander
 Hispanic White Non-Hispanic
 Other Black Non-Hispanic
7. Is English your first language?
 Yes No
8. About how many hours per week do you spend viewing a computer screen? _____

Post Test Questions:

In this box,
please write the line
number of the last word
that you read.

In this box,
please write the last two
words that you read.

Comprehension Test Answers:

- | | | | |
|-------------|-------------|-------------|-------------|
| 1 (●●●●●●) | 11 (●●●●●●) | 21 (●●●●●●) | 31 (●●●●●●) |
| 2 (●●●●●●) | 12 (●●●●●●) | 22 (●●●●●●) | 32 (●●●●●●) |
| 3 (●●●●●●) | 13 (●●●●●●) | 23 (●●●●●●) | 33 (●●●●●●) |
| 4 (●●●●●●) | 14 (●●●●●●) | 24 (●●●●●●) | 34 (●●●●●●) |
| 5 (●●●●●●) | 15 (●●●●●●) | 25 (●●●●●●) | 35 (●●●●●●) |
| 6 (●●●●●●) | 16 (●●●●●●) | 26 (●●●●●●) | 36 (●●●●●●) |
| 7 (●●●●●●) | 17 (●●●●●●) | 27 (●●●●●●) | 37 (●●●●●●) |
| 8 (●●●●●●) | 18 (●●●●●●) | 28 (●●●●●●) | 38 (●●●●●●) |
| 9 (●●●●●●) | 19 (●●●●●●) | 29 (●●●●●●) | |
| 10 (●●●●●●) | 20 (●●●●●●) | 30 (●●●●●●) | |

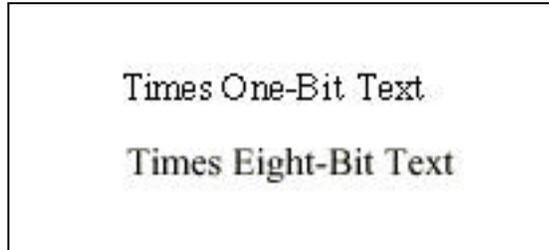
After you've completed the test, please answer the following questions:

1. Could you easily distinguish the letter "e" from the letter "o"? Yes No
2. Could you easily distinguish the number "5" from the letter "S"? Yes No
3. Could you easily distinguish the letter B from the number "8"? Yes No
4. Could you easily distinguish the letter "c" from the letter "e"? Yes No
5. Could you easily distinguish the letter "o" from the letter "c"? Yes No

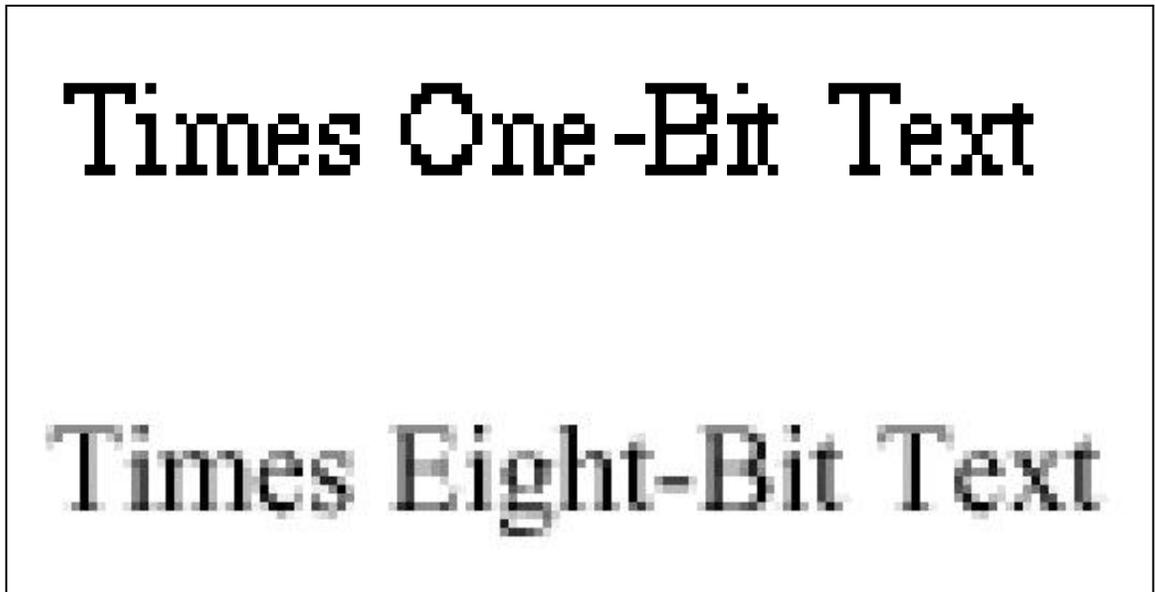
Thank you for your participation in this study!

APPENDIX E:

Times Displayed as 1-bit and 8-bit Text



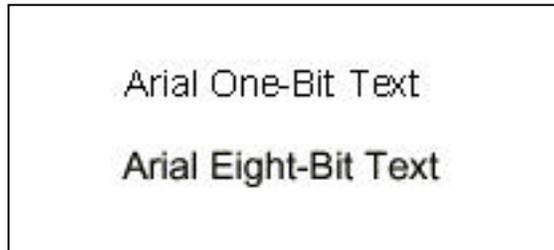
(100% Magnification)



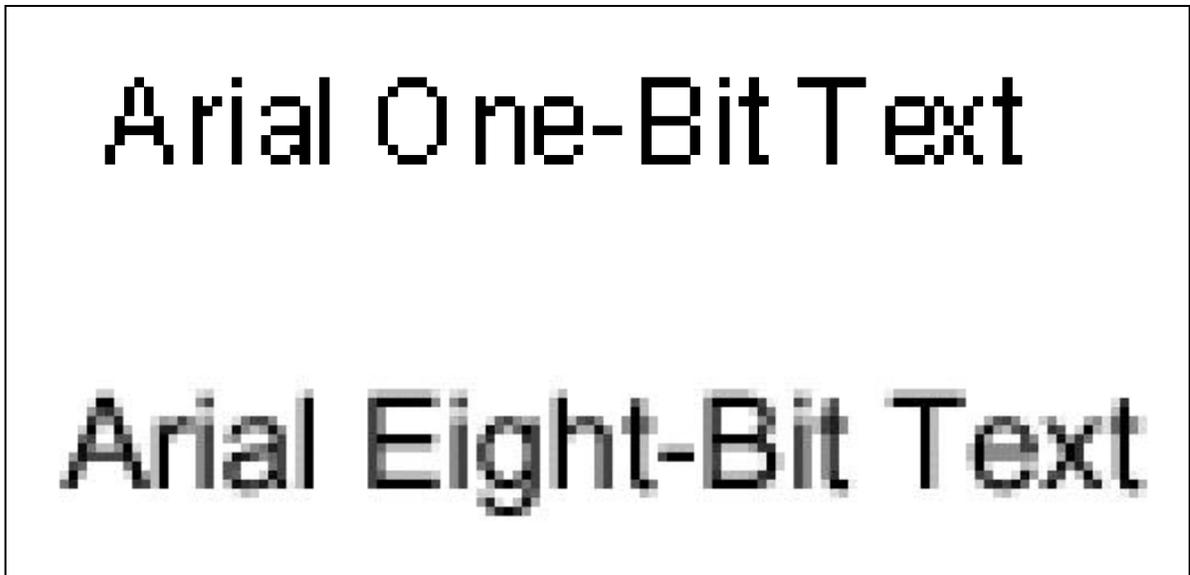
(400% Magnification)

APPENDIX F:

Arial Displayed as 1-bit and 8-bit Text



(100% Magnification)



(400% Magnification)

APPENDIX G:

Georgia Displayed as 1-bit and 8-bit Text

Georgia One-Bit Text
Georgia Eight-Bit Text

(100% Magnification)

Georgia One-Bit Text
Georgia Eight-Bit Text

(400% Magnification)

APPENDIX H:

Verdana Displayed as 1-bit and 8-bit Text

Verdana One-Bit Text
Verdana Eight-Bit Text

(100% Magnification)

Verdana One-Bit Text
Verdana Eight-Bit Text

(400% Magnification)

Scholarly Productivity:

Research

Weisenmiller, E.M. (1999). *A Study of the Readability of On-Screen Text*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg.

Weisenmiller, E.M. (1995). *The Impact of the Macintosh PowerPC on the Prepress Industry of the Southeastern United States*. Unpublished masters thesis, Georgia Southern University, Statesboro.

Publications

Weisenmiller, Eric (1996). The Impact of the Macintosh PowerPC on the Prepress Industry of the Southeastern United States. *Spotting the News: Official Publication of the Southeast Prepress Association* .50(5). 5-8.

Presentations:

Weisenmiller, E. M. (February, 1996). *Using HTML to Create World Wide Web Pages*. Southeastern Prepress Association Technical Update, Atlanta, Georgia.

Weisenmiller, E. M. (September, 1997). *Graphic Comm Central*. GATF Teacher Conference at Print'97, Chicago, Illinois.

Service Activities:

Memberships

- Epsilon Pi Tau
- International Graphic Arts Education Association
- NAITTE

Committees

- Homecoming Committee (1998-1999)
- Awards Committee (1998-1999)
- Departmental Faculty Status Committee (1999-2001)

VITA

Eric Michael Weisenmiller

Degrees:

Ph.D., Curriculum & Instruction, Virginia Polytechnic Institute & State University, 1999

M.T., Industrial Technology, Georgia Southern University, 1995

B.A., English, Georgia Southern University, 1993

Professional Experience:

*August 1998 -
Present*

Assistant Professor
Illinois State University, Normal, Illinois

*August 1995 -
May 1998*

Graduate Teaching Assistant
Virginia Polytechnic Institute and State University,
Blacksburg, Virginia,

*September 1993 -
August 1995*

Graduate Assistant
Georgia Southern University, Statesboro, Georgia

Areas of Specialization and Courses Taught:

Specialization: Desktop Publishing
Electronic Imaging and Information Systems
Prepress Management
Technology Education
Curriculum Development

Courses Taught: IT 150 Fundamentals of Printing & Imaging Technology
IT 151 Introduction to Industrial Computing Systems
IT 250 Electronic Imaging Technologies
IT 253 Photomechanical & Digital Imaging Systems
IT 283 Information & Imaging Systems
IT 352 Advanced Electronic Page Composition

Other Activities

- Judge, Newsletter Design Competition, Phi Delta Kappa (Southeastern Region)
- Assisted in the maintenance of the official web site (<http://www.igaea.org>) of the International Graphic Arts Educators Association (IGAEA).
- Took part in a team project to redesign the web site for Holy Cross Catholic Church, Lynchburg, Virginia (<http://www.lynchburg.net/community/churches/holycross/>).
- Copy edited 1994 edition of the Georgia Southern University Printing Management Annual Report.
- Art Director, Southern Reflector, Georgia Southern University's quarterly magazine (September, 1992 - June 1993). Supervised writers, computer composition staff, and photography staff, and oversaw Macintosh prepress production system.
- Managing Editor, The George-Anne, Georgia southern's student-run newspaper (June, 1991 - June 1992). This position required intensive involvement in the daily operation of the newspaper, direct oversight of the editorial and production phases of the twice-weekly publication. During my tenure, I was successful in converting the production process of the paper from a manual process to a Macintosh digital desktop prepress production system.
- Staff Writer, 1992 NCAA Division I-AA Football Championship Special Edition of the George-Anne, Georgia Southern University's Official Student Newspaper sponsored by the National Collegiate Athletic Association.
- Assisted in the development of GRAPHIC COMM CENTRAL (<http://Teched.vt.edu/gcc>), a WWW site that is intended to provide a wide range of information and services for Graphic Communication educators, students, and industry personnel.