DOCUMENT RESUME

ED 397 784 IR 017 972

AUTHOR Chen, Ming-Pu; And Others

TITLE The Effects of Font Size in a Hypertext Computer Based Instruction Environment.

PUB DATE 96

NOTE 9p.; In: Proceedings of Selected Research and Development Presentations at the 1996 National Convention of the Association for Educational Communications and Technology (18th, Indianapolis, IN, 1996); see IR 017 960.

PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS College Students; Computer Assisted Instruction; Computer Science Education; Computer System Design; *Design Preferences; Higher Education; Instructional Materials; Material Development; Reading Comprehension; *Screen Design (Computers); Student Attitudes; *User Satisfaction (Information)

IDENTIFIERS *Legibility; *Typography

ABSTRACT

Some of the elements instructional designers must consider when designing screens for computer based instruction (CBI) include legibility, readability, and particularly font size. The purpose of this study was to determine which size font is most effective for CBI. Participants were 163 undergraduate and graduate students enrolled in an upper-level computer course at a western university. Subjects were randomly assigned to one of four treatment groups. Each treatment group received a different font size. All subjects received a computerized version of the Nelson-Denny Reading Test Form E. The four dependent measures were reading comprehension, satisfaction, preference, and distance from computer screen. The three independent measures were font size, age, and section. Results indicated a lack of satisfaction with the assigned font sizes of 10 and 12 points and satisfaction with font sizes of 14 and 16 points. Subjects preferred the 16-point font size to other sizes. (Contains 24 references.) (AEF)

Title:

The Effects of Font Size in a Hypertext Computer Based Instruction Environment

Authors:

**Ming-Pu Chen, Walter A. Jackson, Clark Parsons,
Kathleen M. Sindt, Jennifer B. Summerville, Donald D. Tharp, Ronald R. Ullrich, Edward P. Caffarella**
EPSAT
College of Education
University of Northern Colorado

Some of the elements instructional designers must consider when designing screens for computer based instruction (CBI) include legibility, readability, and particularly font size. These elements are also factors that influence the comprehension and speed with which people learn (Grabinger, 1992). Although Grabinger proposes that no single element promotes learning more than another; the combination of these elements on the computer screen affects learning more than any single element. The contribution of font size in the development of CBI deserves investigation. The purpose of this study was to determine which size font is most effective for computer based instruction users.

Several research studies were conducted between 1885 and 1963 on the legibility of print based information. Tinker (1963) reported that no single method of measurement is adequate for determining the legibility of print in all kinds of typographical setups. Tinker provides several guidelines, based on the research prior to 1963, for the optimal legibility in print materials. These guidelines include (a) leading (space between lines) has an important effect on the legibility of type, especially in fonts smaller than 12 point and (b) font sizes between 8 and 12 points are acceptable. Eleven point font is the best size for reading speed and comprehension.

Many of the recommendations offered in the literature lack research support. One example is the recommendation of Felker, Pickering, Charrow, and Holland (1981) of 8 to 10 point font for the printed page of documents, and 14 point font for headings on documents. These authors reported the most effective font size should be easy to read and pleasing to the eye but offered no real evidence of what constituted ease of reading and what is pleasant to the eye.

A review of literature revealed very little direct research on computer font size. Similarly there was only limited research concerning font size for television screens. A television screen is an interface similar to a computer monitor but there are some major differences. These differences are television is not typically used for extensive reading, the viewer to screen distance is usually greater than the typical distance of one to three feet for a computer screen, and the resolution and the physics of screen refresh for a television screen is different from a computer screen.

Although the television research results do not apply directly to this study, some of the techniques and concepts from this research do apply. One of these is the more general "subtended angle" approach rather than trying to identify specific fonts for specific distances. Nevertheless, this study had a different medium and different tasks to study

With the advent of television, researchers began studying the legibility of characters on television screens. Smith (1979) showed a proportional width font is significantly better for recognition of words on TV screens than a non-proportional font. Research shows that character size ranging from 10 to 24 minutes of vertical arc is best for legibility on television screens (Shurtleff, 1967). Minutes of arc refer to the angle defined by the top of the character, the eye, and the bottom of the character. This angle is the subtended angle.

Shurtleff (1967) reported on legibility and symbol size on television screen in a review of literature that gave guidelines for accuracy of reading to occur. The examination found that viewers required a visual size of between 12 and 15 minutes of arc and vertical resolution between 8 and 12 scanning lines per symbol height.

Legibility involves perceiving letters and words, and the reading of continuous textual material. In the final analysis, optimal legibility of print is achieved by typographical arrangement in which shape of letters and other symbols, characteristic word forms, and all other typographical factors such as type size, line width, leading, page layout, and density are coordinated to produce comfortable vision, and easy rapid reading with comprehension.

Computer technology is rapidly becoming a component in training and education. People are using computers to learn new concepts and reinforce previously learned cognitive and motor skills. One such issue of interest is the legibility of the text on the computer screen. Legibility affects several aspects of learning, including the reading speed of the user and the comprehension of the information by the user (Tinker, 1963). Therefore, the legibility of the text on the computer screen is an item of importance to designers of computer based instruction.

Several authors (Heines, 1984; Hooper & Hannafin, 1986; Criswell, 1989) provide guidelines for optimal font size, leading, blank space, and line justification on computer screens. These guidelines are similar across the literature and are as follows: font size for computer screens should be 10 or 12 point, leading should be 15% of the character width, and left justification should be used, leaving the right edge of the screen jagged (Brockman, 1990). There is some difference in opinion about the recommended length of the line, with Hooper and Hannafin (1986) recommending up to 80 characters per line (the entire screen width) and Brockman (1990) and Merrill (1982) recommending that white space be left on the edges of the screen. While these guidelines are useful for designers of computer based instruction, there is little research using computer screens to support these guidelines. This study investigated the relationship of font size, comprehension, and user preference.

One aspect of legibility is the font size of the text displayed. While there is research relating to font size for printed information, little is available on font size as related to legibility when text is displayed on the computer screen. There are guidelines provided for use of font size on computer screens, but these guidelines tend to be based on print research (Hooper & Hannafin, 1986).

Giddings (1972) conducted a legibility experiment in an effort to determine optimum height for characters under particular conditions. Giddings' experiment attempted to find suitable sizes for words and numbers appearing on a directly viewed CRT. The results indicated that differences were found in the legibility of 6 letter words and single digits. Giddings reported that this may be due more to the informational content of the words and digits than their fundamental nature.

Snyder & Taylor (1979) conducted similar research concerning: font size and dot luminance of alphanumeric characters in comparison to recognition accuracy, response time, tachistoscopic recognition, and threshold visibility when presented in a non-contextual form on a CRT display. Tachistoscopic recognition refers to the length of time a display remains on a viewer screen. Threshold visibility pertains to the distance at which a viewer can recognize characters on a viewer screen. Findings indicated accuracy improved as font size increased and regressed as distance increased. The effect of higher levels of luminance was greater accuracy for the smaller character sizes and the further distances from the display screen. Response time increased as character size decreased and viewing distance increased.

Further, increases in luminance improved tachistoscopic recognition accuracy more for the smaller character sizes than larger. No significant findings were produced for threshold visibility. The conclusion by Snyder & Taylor (1979) was that characters less than 4.79 mm in height at a viewing distance of 1.5 m will significantly reduce legibility.

Similar research by Beldie, Pastoor, & Schwarz (1983) examined the question: Are variable-matrix characters more legible than fixed-matrix letters on readability, error identification, and the ability to locate the position of the target word on a CRT screen? In this experiment, fixed width characters occupied the same amount of horizontal space. Variable width proportional characters occupied differing width spaces depending on the character. Variable width characters such as "i" occupied less horizontal space than "w." The findings suggested that a proportional width font, even without serifs is significantly better for recognition of words on television screens. The authors recommended it for text displays.

Hathaway's (1984) literature review revealed interesting findings regarding letter size and density of display text. In a review of Smith's (1979) research, Hathaway examined letter size as it interfaced with legibility. Smith found that the viewers sitting 24 inches from a computer screen would need letters no larger than 3/16", and 94 percent of viewers would need letters no larger than 3/32" high from the same distance.

Readability

Comprehension and readability are important factors when designing screen layout and selecting the most effective size of font. Presenting lessons effectively on the screen is one of the central problems in computer based learning (Isaacs, 1987). Although font size is an important issue when designing computer screens, the instructional designer must consider typographic factors, organization factors, cueing factors, and control factors, in addition to the content (Grabinger, 1989).

Density of displayed text on a CRT impacts the speed with which readers comprehend material (Kolers, 1981; Jonassen, 1982; Hooper & Hannafin, 1986). Subjects read faster and more accurately with double spaced text than with single spaced text. Koler's findings indicated 80 characters per line resulted in faster and more accurate reading. Both the Jonassen, and the Hooper & Hannafin studies suggested that ragged right margins are easier to read than right justified lines.

Milheim & Lavix (1992) suggest using various type sizes and fonts to emphasize certain materials and variety. These suggestions include using bold type face for higher-level information such as titles or headings and using mixed-case letters for higher legibility and faster reading. Further, Bailey and Milheim (1991) recommend using only one font style per screen, text sizes larger than 12 point, and primarily using New York or Geneva fonts. Faiola (1990) suggests using screens that are consistent in size, color, and fonts for text, since more than two or three typefaces and/or sizes used within one screen can be disconcerting to learners.

Grabinger (1988, 1992) studied the effects of single and multiple dimensions on the computer screen in relation to learners' ability to read a computer screen (readability) and the understanding and learning that followed the viewing of the screen (study-ability). The purpose of this research was to identify constructs that could guide the design of computer screens used to display information in computer-assisted instruction, hypermedia, or on-line help applications. Grabinger reported comprehensibility is the combination of readability and study-ability. He further reported that single element research is extremely important in identifying the strengths, weaknesses, and potential problems of using specific attributes on CRT screens, such as font size.

Grabinger's findings indicate that screens that are organized according to functional areas with space, boxes, and lines, the use of headings, directive cues, and spaced paragraphs combined to present planned, controlled, organized, and structured appearances are the best for learning conceptual ideas. Further, visual interest is an important element. Grabinger reported that screens that are plain, simple, unbalanced, and bare are perceived as undesirable, while a moderate degree of complexity creates an environment that invited further exploration, in other words, motivation to continue reading.

On the other hand, in related research, Morrison, Ross, & O'Dell (1988) suggest that low density narratives promote better learning and more favorable attitudes on CBI lessons by reducing reading and cognitive processing demand of screen displays. The notion here is instructional designers, at the time of publishing, use the same design formats and teaching strategies traditionally incorporated in print lessons for developing CBI lessons. The rationale behind the study is that computers impose constraints on the display of instructional text because they offer considerably less flexibility than books. Computers limit the visible display to one page at a time, make backward paging for review more difficult, and limit the size of the page layout to 24 lines and 40-80 characters.

Pastoor, Schwarz, & Beldie (1983) conducted research on font sizes presented on standard home color television sets. Subjects performed five different tasks and rated the various character sizes. Their qualitative performance for all tasks was the same with each of the different character sizes; the amount of time necessary to complete the task, however, varied up to more than 20%. This would seem to indicate that none of the character sizes exhibited deficiencies that would exclude their use on color television screens. However, considering the time variance measured, and the fact that the larger font sizes decreased the time with which subjects could perform the tasks, the larger fonts appeared to be better for use on television screens. In all tasks, smaller character sizes produced poorer performance.

Research Questions

The purpose of this study was to find a font size, or sizes, to recommend for presenting text on a computer monitor. There are some obvious limits on the font size because of the discriminatory limits of the eye on the inability of the pixel resolution of a monitor to differentiate one character from another at extremely small sizes. There is also a limitation at which font sizes are too large to read. In the extreme, a 1000 point font would be almost 15 inches tall, much too tall to see even one letter on most computer monitors. Somewhere in between these extremes are the appropriate font sizes that users normally see on computer screens, such as 12 point, 16 point, and 24 point. The assumption is there is some determinable range of these font sizes that are readable, and a subset of the range that especially facilitates reading speed and comprehension. Finding the appropriate subset was the goal of this research.

Another question of interest is whether users might prefer one font size over another when reading for speed and comprehension. If there are several font sizes that promote optimal reading speed and comprehension, user preference might be a factor to help direct a font choice.

Methodology

Participants

The sample consisted of one hundred and sixty-three undergraduate and graduate students enrolled in an upper-level computer course entitled "Classroom Applications of Educational Technology" at a western university. The age of the volunteer subjects ranged from 20 to 46.

Subjects were randomly assigned to one of four treatment groups. Each treatment group received a different font size. Group one with 47 subjects used 10 point font. Group two with 37 subjects used 12 point font. Group three with 47 subjects used 14 point font. Group four with 32 subjects used 16 point font.

All subjects received a computerized version of the comprehensive section of the Neson-Denny Reading Test Form E (1991). The comprehensive section of the Nelson-Denny Reading Test consists of eight passages with one to eight paragraphs in each followed by four to eight questions. The passages were presented in a scrolling window on the screen. When the student completed reading the passage, the student moved to another screen and answered the questions on the passage. The student could not go back to the passage and review. This procedure was repeated for all eight passages.

The passages were presented on Apple Macintosh Centris model 610 and 650 computers with 14 inch display monitors. The passages were displayed with mixed case black letters on a white background in a Times Truetype font style. The font sizes were 10, 12, 14, or 16 point depending on the treatment group.

Procedure

The subjects received the following instructions before reading the first passages: (a) position themselves to a height with their eyes between the middle of the screen and the top of the screen, (b) seat themselves between 24" and 30" from the screen, and (c) enter their year of birth and await further instructions. After each student completed these tasks they were instructed about the passage content, number of questions, and the 15 minute time limit.

The subjects proceeded through the reading comprehension test at their own pace. The size of the text on the screen was randomly assigned by the computer program. Upon completion of the test, the subjects responded to a question about font size satisfaction. If unsatisfied, the subjects chose a font size from 10, 12, 14, and 16 point size.

The subjects were observed to determine those individuals not within the range of 24" to 30" from the screen. These subjects who were not within the range were noted in the experimental log.

Results

Dependent measures

There are four dependent measures: reading comprehension, satisfaction, preference, and distance from computer screen. Reading comprehension measured participants' reading comprehension of test passages. Satisfaction measured participants' satisfaction with assigned font size. Preference measured participants' preferred font size after the treatment. Posture indicates whether the students were inside or outside the optimal 2-2 1/2 feet from the screen during the reading of the passages.

Independent Measures

There are three independent measures: font size, age, and section. Subjects were randomly assigned to the font size groups by the computer program and were assigned sections based on which class they were in.

Design and data analysis

An ANCOVA was conducted using the dependent variable reading comprehension and the independent variables of font size and section. No reading comprehension differences were found for the main effects (F=0.35, p=.78 for font size and F=0.61,p=.77 for section) or interaction (F=1.23, p=.23) at a = 0.05. Likewise, the covariance of age did not effect the reading comprehension (F=0.13, p=.71). The ANCOVA source table is listed in Table 1.

Table 1
ANCOVA for Font Size by Section with Covariate of Age

Source	DF	Sum Of Squares	Mean Square	F
Font Size	3	21.64	7.21	0.35
Section	8	98.70	12.34	0.61
Font Size by Section	21	524.68	24.98	1.23
Age	1	2.67	2.67	0.13
Error	127	2584.52	20.35	
Total	160	3297.98		

A Chi-square was performed to examine whether a significant difference occurred between font size and satisfaction. A significant difference, shown in Table 2 (Chi-square = 74.204, p <.001) was found for satisfaction at a =.05.

Table 2
Chi-Square of Font Size by Satisfaction

	Satisfaction			
Font Size	No	Yes	Total	Chi-square
10	44	1	45	74.204
	97.78	2.22		
12	31	6	37	
	83.78	16.22		
14	22	25	47	
	46.81	53.19		
16	3	29	32	
	9.38	90.63		
Total	100	61	161	
	62.11	37.89	100.00	

A Chi-square was performed to examine whether a significant difference occurred between font size and preference (SeeTable 3). A significant difference (Chi-square = 51.81, p <.001) was found for preference at a =.05.

Table 3
Chi-square of Font Size by Preference

Font Size	Preference 10	12	14	16	Total	Chi-square
10	1	0	24	20	45	51.81
	2.22	0.00	53.33	44.44		
12	0	6	14	17	37	
	0.00	16.22	37.84	45.95		
14	0	0	27	20	47	
	0.00	0.00	57.45	42.55		
16	0	0	1	31	32	
	0.00	0.00	3.13	96.88		
Total	1	6	66	88	161	
	0.62	3.73	40.99	54.66	100.00	

A Chi-square was performed to examine whether a significant difference occurred between subject's distance from screen and font size. No significant difference (Chi-square = 4.78, p = 0.573) was found at a =.05. See Table 4 for the Chi-square results.

Table 4
Chi-square of font size by distance from screen

Font Size	Distance Leaning Forward	Normal	Leaning Back	Total	Chi-square
10	14	31	0	45	4.78
	31.11	68.89	0.00		
12	8	28	1	37	
	21.62	75.68	2.70		
14	16	29	2	47	
	34.04	61.70	4.26		
16	10	22	0	32	
	31.25	68.75	0.00		
Total	48	110	3	161	
	29.81	68.32	1.86	100.00	

Discussion

The research hypothesis that an optimal font size for CBI design could be determined by our efforts lacked support. Reading comprehension scores did not show a significant difference. Most subjects needed more than the allowed fifteen minutes to complete the passages.

Results indicated a lack of satisfaction with the assigned font sizes of 10 and 12 points, however they were satisfied with font sizes of 14 and 16 points. Subjects preferred the 16 point size to the other font sizes. This result indicates a possible ceiling effect for font size at 16 points. From this we may infer that optimal font size recommendations of 10 - 12 point derived from print media research (Hooper & Hannafin, 1986; Tinker, 1963) are not adequate for CBI applications. Further our results conflict with those of Heines (1984) and Criswell (1989) which recommend a font size of 10 to 12 points. However, our results do agree with those of Pastoor, Schwarz & Beldie (1983) and Bailey and Milheim (1991) which favor larger font sizes.

This study had a number of limitations. The first group of subjects all received the 10 point font. This unexpected result apparently occurred as a consequence of the way HyperCard began its random number generation function. Subjects could have been randomly assigned in advance and not randomly assigned by group from the HyperCard random number generator. Legibility could have been more closely controlled. The reading test was expected to be completed within 15 minutes. However, only 8% of the subjects completed all passages and questions. Subjects should have been allowed more time to complete the passages and questions or fewer passages and questions should have been assigned. This study did not take into account other font styles or the use of double spacing.

The results of this study have two primary implications. Future research is needed because font size recommendations based on print media are inadequate, and the optimal size font for CBI applications is still not known. Future research should include the use of larger font sizes and take the above limitations into account.

Bibliography

Bailey, H. J., Milheim, W. D., (1991). A comprehensive model for designing interactive video based materials. Proceedings of the Ninth Conference on Interactive Instruction Delivery at the 1991 Society for Applied Learning Technology Conference. (February), Orlando, FL.

Beldie, I. P., Pastoor, S., & Schwarz, E. (1983). Fixed versus variable letter width for televised text. Human Factors, 25 (3), 273-277.

Brockman, R. J. (1990). Writing better computer user documentation: From paper to hypertext. (2nd ed.) New York: John Wiley & Sons, Inc.

Criswell, E. L. (1989). The Design of Computer-Based Instruction. New York, Macmillan Publishing Company.

Faiola, T. (1990). Principles and guidelines for a screen display interface. The Videodiscs Monitor, 8 (2), 27-29.

Felker, D. B., Pickering, F., Charrow, V. R., Holland, V. M. & Redish, J. C. (1981). Guidelines for Document Designers. Washington DC: American Institutes for Research.

Giddings, B. J. (1972). Alphanumerics for raster display. Ergonomics, 15, 65-72.

Grabinger, R. S. & Amedeo, D. (1988). CRT Text layout: Perceptions of viewers. Journal of Computers and Human Behavior, 4, 189-205.

Grabinger, R. S. (1989). Screen layout design: Research into the overall appearance of the screen. Computers in Human Behavior, 5 175-183.

Grabinger, R. S. (1992). Computer screen designs: Viewer judgments. Educational Technology Research & Design, 41 (2), 35-73.

Hathaway, M. D. (1984). Variables of computer screen display and how they affect learning. Educational Technology, 24 (1), 7-11.

Heines, J. M. (1984). Screen design strategies for computer-assisted instruction. Washington, DC: Association for Educational Communications and Technology.

Hooper, S. & Hannafin, M. J. (1986). Variables affecting the legibility of computer-generated text. Journal of Instructional Design, 9 (4), 22-28.

Isaacs, G. (1987). Text screen design for computer-assisted learning. British Journal of Educational Technology, 18 (1), 41-51.

Jonassen, D. H. (Ed.) (1982). The Technology of Text. Englewood Cliffs, NJ: Educational Technology Publications.

Kolers, P. A., Duchnicky, R. L., Ferguson, D., C. (1981). Eye movement measurement of readability of CRT displays. Human Factors, 23, 517-528.

Merrill, P. (1982). Displaying text on microcomputers. In D. Jonassen (Ed.), The Technology of Text. (Vol. 1, pp. 401-414). Englewood Cliffs, NJ: Educational Technology Publications.

Milheim, W., Lavix, C. (1992). Screen designs for computer-based training and interactive video: Practical suggestions and overall guidelines. Performance and Instruction, (May/June), 13-21.

Morrison, G. R., Ross, S. M., & O'Dell, J. K. (1988). Text density level as a design variable in instructional display. Educational Communication & Technology Journal, 36 (1), 103-115.

Pastoor, S., Schwarz, E., & Beldie, I. P. (1983). The relative suitability of four dot-matrix sizes for text presentation on color television screens. Human Factors, 25 (3), 265-272.

Shurtleff, D. A. (1967). Studies in television legibility - a review of the literature. Information Display, 4 (1), 40-45.

Smith, S. L. (1979). Letter size and legibility. Human Factors, 21 (6), 661-670.

Snyder, H. L., & Taylor, G. B. (1979). The sensitivity of response measures of alphanumeric legibility to variations in dot matrix display parameters. Human Factors, 21 (4), 457-471.

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