Admin

UMTRI-88-34

78865

LEGIBILITY OF TEXT ON INSTRUMENT PANELS: A LITERATURE REVIEW

Paul Green
Steve Goldstein
Kristine Zeltner
Susan Adams

NOVEMBER 1988

UMTRI The University of Michigan
Transportation Research Institute

Technical Report Documentation Page

1. Report No. UMTRI-88-34	**2. Government Accession No.**	**3. Recipient's Catalog No.**

4. Title and Subtitle

LEGIBILITY OF TEXT ON INSTRUMENT PANELS:
A LITERATURE REVIEW

5. Report Date November, 1988

6. Performing Organization Code 389214

7. Author(s) Paul Green, Steve Goldstein, Kristine Zeltner, and Susan Adams

8. Performing Organization Report No. UMTRI-88-34

9. Performing Organization Name and Address

The University of Michigan
Transportation Research Institute
2901 Baxter Road
Ann Arbor, MI 48109-2150 U.S.A.

10. Work Unit No. (TRAIS)

11. Contract or Grant No. DRDA 86-1909-P1

12. Sponsoring Agency Name and Address

Chrysler Motors Corporation
R&D Programs Administration
12000 Chrysler Drive
Highland Park, MI 48288-1118

13. Type of Report and Period Covered January 1, 1987- August 31, 1988 (Final Report)

14. Sponsoring Agency Code

15. Supplementary Notes

Supported by the Chrysler Challenge Fund

16. Abstract

This report reviews 46 documents from the literature pertaining to the legibility of text on instrument panels. The review examines human factors issues for both continuous stroke, multiple segment, and dot matrix characters. Basic human visual performance--the effects of luminance contrast, illumination levels, color, task, and viewer visual acuity on the legibility of simple targets--is covered. Also covered are studies of font, generic models of text legibility, and research on three applications--highway signs, displays in aircraft cockpits, and automotive displays.

The review identifies over a half dozen procedures for calculating legibility (e.g., Peters and Adams, 1959; Duncan and Konz, 1976; Howett, 1983; etc.). Of these, the Bond Rule should be followed when a quick answer is needed. (Character Height=.007 x Viewing Distance.) Otherwise, use the procedure developed in this report based the Mourant and Langolf, 1976 data. According to that expression: Response Time (secs) = 5.82 - 13.03H -.70Log(L) + 2.94/C where H is the Character Height (in), L the Character Luminance (ft-L), and C is the Contrast Ratio.

17. Key Words

Human factors, ergonomics, human engineering, displays, legibility, instrument panels, automobiles, cars, engineering psychology

18. Distribution Statement

19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 164	**22. Price**

Reproduction of completed page authorized

TABLE OF CONTENTS

- *Table of Contents* -

- *Table of Contents* -

LIST OF TABLES

- LIST OF TABLES -

LIST OF FIGURES

- LIST OF FIGURES -

- *LIST OF FIGURES* -

- LIST OF FIGURES -

EXECUTIVE SUMMARY

Green, P., Goldstein, S., Zeltner, K, and Adams, S. (1988). Legibility of Text on Instrument Panels: A Literature Review (Technical Report UMTRI-88-34). Ann Arbor, MI: The University of Michigan Transportation Research Institute.

This report concerns the first task in a four-task project entitled "Recognition and Comprehension of Electronic Display Graphics." This research was supported by the Chrysler Corporation Challenge Fund. This project provides information designers and engineers can use to make displays that will be legible, understandable, and, consequently, easy to use.

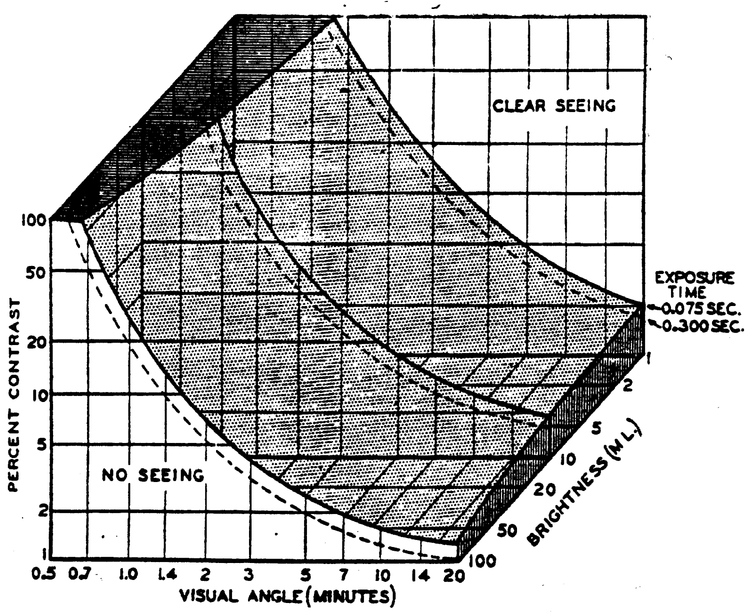
This is the third of three reports reviewing the literature on the legibility of text (Task 1). This report addresses five questions:

1. *How do fundamental lighting variables (luminance contrast and illumination level) and exposure duration affect people's ability to detect simple visual targets?*
2. *What is the effect of chromatic contrast on legibility?*
3. *What is the effect of font on the legibility of text?*
4. *How well do adults see?*
5. *What expressions are there to predict the legibility of text for various applications?*

1. How do fundamental lighting variables (luminance contrast and illumination level) and exposure duration affect how well people detect simple visual targets?

The data most commonly referred to on legibility thresholds are the Cobb and Moss curves (as replotted by Luckiesh and Moss, 1937), which follow. They show that the most important variable affecting legibility is the contrast ratio, with illumination levels and exposure duration having secondary effects. Other studies show that legibility thresholds depend on the visual angle of the target, but are fairly independent of distance.

Predictions for legibility thresholds for a variety of simple targets have also been developed by Moon and Spencer (1944) and, in a series of studies, by Blackwell. Blackwell's work is described in detail in the report. Procedures for using those results to compute the legibility of text are also given.



Luckiesh and Moss (1937) Version of Cobb and Moss Curves.

2. What is the effect of chromatic contrast on legibility?

In general, the effects of chromatic contrast on legibility are relatively small. Two expressions of chromatic contrast appear in the literature, one based on the CIELAB data (CIE Yu'v') and the other based on CIELUV. They are listed below in that order.

$$\text{Delta\_E} = [\ (\text{Delta\_L}^*)^2 + (\text{Delta\_u}^*)^2 + (\text{Delta\_v}^*)^2\ ]^{0.5}$$

where:
$L^* = 116(Y/Y0)\hat{}1/3 - 16,\ Y/Y0>.01$
$u^* = 13L^*(u'-u0')$
$v^* = 13L^*(v'-v0')$
$u' = 4X/(X + 15Y + 3Z)$
$v' = 9Y/(X + 15Y + 3Z)$

$$\text{Delta\_E}(Yu'v') = [(155\ TB/M)^2 + (367\ U)^2 + (167\ V)^2]^{0.5}$$

where: TB = difference in luminance between text and background
M = maximum luminance of text or background
U = difference between text and background u' coordinates
V = difference between text and background v' coordinates

The ANSI standard for office workstations requires that characters presented on screens have Delta_E values in excess of 100 (CIE Lu*v*).

3. What is the effect of font on the legibility of text?

The literature shows that the performance differences between fonts are relatively small. However, fonts should not be ignored, as font modifications are quite straightforward.

Beyond that, the literature shows that confusion errors are very predictable for seven-segment displays. The frequency of particular confusions depends upon the number of segments by which the character pair in question differ. For limited character sets (e.g., numbers only) Van Nes and Bouma offer design suggestions to minimize reading errors.

With regard to dot matrix characters, the key studies are those of Snyder and Maddox (1978) and Shurtleff's 1980 book. They show that the character matrix should be at least 7x9 for easy reading with 9x11 preferred. Dots should be close together and round.

4. How well do adults see?

If displays are to be easily read, engineers must know how well viewers see. There are two useful studies on this topic. A British study (Davison and Irving, 1980) considers only drivers, while the U.S. work (Roberts, 1964), carried out as part of the national health survey, concerns all adults. Both data sets provide useful information. They clearly show a marked decline in population acuity at age 45, and that the corrected acuity of a large number of older drivers is 20/40 or worse. Tables of data from both reports appear in this report.

5. What expressions are there to predict the legibility of text for various applications?

There are several formulas in the literature that summarize experimental work on the legibility of text on displays. They include:

a. Peters and Adams (1959)

Letter Height (inches) = $H = .0022D + K1 + K2$

where:
D = Viewing Distance (inches)
$K1$ = 0.06 for > 1.0 fc, favorable reading conditions
= 0.16 for > 1.0 fc, unfavorable conditions or < 1.0 fc, favorable conditions
= 0.26 for < 1.0 fc, unfavorable conditions
$K2$ = 0.075 for emergency labels, counters, scales, legend lights
= 0.0 for other (unimportant) panel markings

b. Mourant and Langolf (1976)

Response Time (seconds) = $RT = 5.82 - 13.03H - .70\log(L) + 2.94/C$

where:
H = Height (inches)
L = Character Luminance (foot-Lamberts)
C = Contrast Ratio

c. Duncan and Konz (1976)

$\text{Height (cm)} = H = .0015De + .0519(H{:}Sw) - .3499$

where: De = No Error Viewing Distance (cm)
H:Sw= Height:Strokewidth Ratio

- or -

$H = .0038Dp + .0385(H{:}Sw) - .0864$

where: De = Preferred Viewing Distance (cm)
H:Sw= Height:Strokewidth Ratio.

d. Smith (1978) - (The Bond Rule)

$\text{Height} = .007 * \text{Viewing Distance}$

e. Military Standard 1472C (U.S. Department of Defense, 1981)

	Height	
Marking	<=3.5 cd/m^2 (1 fL)	> 3.5 cd/m^2
critical, variable pos	5-8mm (.2-.31in)	3-5 (.12-.2)
critical, fixed pos	4-8mm (.16-0.31in)	2.5-5 (.1-.2)
non-critical	1.3-5mm (.05-.2in)	1.3-5 (.05-.2)

f. Howett (1983)

.1 $\text{Contrast (\%)} = C = ((Lb - Lt) / Lb) * 100$ (assumes Lb>Lt)

where: Lb = Background Luminance
Lt = Target Luminance

.2 $\text{Snellen Acuity} = S = Sd * (85 / Lb)^{.213} * (90 / C)^{.532}$

where: Sd = denominator in the Snellen ratio.
(e.g., If a viewer has 20/40 visual acuity, use 40.)
Lb = Background Luminance (cd/m2)

.3 $\text{Height} = (H{:}Sw) * 1.45 * 10^{-5} * S * D$

where: H:Sw = Height to Strokewidth Ratio (for 6:1 use 6)
D = Viewing Distance (m)

Of these expressions, the one based on the Mourant and Langolf data is the most appropriate for automobile instrument panel design. It is the only expression listed here in which visual search was part of the test conditions. For simple problems, the Bond Rule should be considered. The report also describes predictive data and models from Payne (1983), Rogers, Spiker, and Cicinelli (1986), and Sawyer and Talley (1987).

PREFACE

This report describes the results of the first task in a four-task project entitled "Recognition and Comprehension of Electronic Display Graphics." The goal of this project is to provide information that designers and engineers can use to develop legible and understandable automotive displays. In particular, this report was intended to provide background information to guide the other phases of this project. Specifically, it identified which factors should be manipulated in the studies of instrument panel legibility and, once the legibility data were collected, what form the model of legibility based on that data should take.

This research was funded by the Chrysler Corporation through the Chrysler Challenge Fund. The purpose of the Fund is to establish closer ties between the Chrysler Corporation and leading American universities, and to promote direct access to the advanced technologies being developed in universities. It also aims to increase interaction between the Chrysler engineering staff and university research personnel, and to increase undergraduate and graduate student awareness of the engineering opportunities available at the Chrysler Corporation.

Other work sponsored by this project includes two additional reviews of the literature on display legibility, several experiments concerned with alternative methods for evaluating legibility, an experiment on the legibility of seven-segment numeric displays, and a review of the literature on human factors in gauge design.

We would like to thank Cathy Colosimo of the Chrysler Corporation for serving as the liaison for this project. Her patience and understanding were greatly appreciated. We would also like to thank Tom Dunn for his insight. Finally, we would like to thank Jim Geschke who was the initial contact person in 1984 when we approached Chrysler about this research and, who saw that it was included in the Challenge Fund Program.

INTRODUCTION

This report is the third of three reports reviewing the literature on legibility as part of the Chrysler Challenge Fund Project on Recognition and Comprehension of Electronic Display Graphics. The first review, Legibility Abstracts from the UMTRI Library (Adams, Goldstein, Zeltner, Ratanaproeksa, Green, 1988) contains references and abstracts for all documents relating to legibility in the University of Michigan Transportation Research Institute (UMTRI) Library. This first review provided information as to the scope and nature of the available legibility literature in the UMTRI Library.

The second review, Selected Abstracts and Reviews of the Legibility Literature (Zeltner, Ratanaproeksa, Goldstein, Adams, and Green, 1988) includes 28 of the 121 documents cited in the first report. It contains revised abstracts emphasizing quantitative and engineering aspects of the research, as well as UMTRI reviews of the research. The central difference between this review and the first review is the inclusion of figures and tables within the revised abstracts.

This third report emphasizes legibility literature from Paul Green's personal files which is not readily available or accessible to the research community. The report concentrates on human factors measures for both continuous stroke and dot matrix displays in terms of character size, color, illumination levels, luminance contrast, task, and viewer visual acuity. However, this report is not an exhaustive review of the legibility literature, but rather concentrates on findings particularly relevant to the Chrysler Challenge Fund Project. Readers interested in other more general reviews concerning legibility should consult the two other literature reviews described above.

This report contains reviews of over 40 documents. These reviews are organized into subsections by topic, and within each topic reviews are arranged chronologically. The report covers four topics with background and basic information appearing first. These topics are: previous literature reviews, factors affecting legibility, adult visual capabilities, and application oriented research. A summary of key findings is also provided.

The material on factors affecting legibility is covered in four sections: basic work on lighting variables, Blackwell's work on that subject, font, and color. The studies examined in these sections are concerned primarily with simple visual targets, such as Landolt C's and vertical bars. A separate section is devoted to Blackwell's work because it is extensive and significant.

The material on applied legibility research is divided into sections on predictions independent of context, studies of highway signs, aircraft applications, and work specifically concerned with automobile instrument panels.

However, it should be noted that this report does not cover material pertaining to how well displays are understood or information concerning the legibility of gauges. Information concerning some of those issues appear in another report associated with this project (Human Factors and Gauge Design: A Literature Review, Green, 1988).

Papers reviewed in this document were selected on the basis of the quality of the work and the relevance to instrument panel display legibility. Some papers which described strong methodology or contained useful general information not necessarily specific to instrument panel displays were also included. The documents reviewed include original research papers, research literature reviews, and papers relating more traditional engineering analyses.

An important difference between this document and traditional literature reviews is the inclusion of figures and tables in the report, and the strong emphasis on quantitative and engineering applications to the problems of legibility. A second major difference is the further analysis done by the authors of this report on some of the documents reviewed. This analysis included additional regression analysis, drawing together the results for a given legibility parameter from several different researchers, and further organization of original data. In many cases, sample legibility calculations are provided to facilitate comparison of the various prediction methods. As a result, this document integrates the findings of several reports, rather than merely being a compilation of the legibility literature. Further, this report also includes summaries for each topic subsection, as well as indices by author and title, and a complete list of references.

PREVIOUS LITERATURE REVIEWS

No literature review would be complete without mention of the previous reviews of the topic. There are several reviews concerning legibility that may be of interest to those involved in automotive design. Those without much knowledge of the topic of human vision may wish to consult a reference before reading further. The authors recommend the classic Chapanis chapter in Applied Experimental Psychology as a possible reference (Chapanis, Garner, and Morgan, 1949). Its title ("How We See") is indicative of the clarity of presentation. A more detailed explanation appears in his chapter in Human Factors in Undersea Warfare (National Research Council, 1949). Another classic is the Wulfeck, Weisz, and Raben (1958) technical report. Several chapters in the Boff, Kaufman, and Thomas (1986) handbook provide a more contemporary perspective. Finally, for those seeking reference information, the IES Handbook (Illuminating Engineering Society, 1972) is recommended.

Following are descriptions of six reviews that are particularly relevant to automotive problems.

Society of Automotive Engineers (1966)

This is the classic publication on basic vision and motor vehicle design. The emphasis of this document is on general tutorial information concerning human performance rather than design recommendations. It consists of two SAE papers, one by Schmidt (SAE 660004) and one by Connolly (SAE 660164), released together.

The Schmidt paper contains a short treatise on physiological optics, shows the standard human dark adaptation function, and threshold functions relating contrast, exposure duration, target size, and accuracy. It also discusses the various measures of visual acuity, presents data on representative levels of illumination, provides equations for calculating glare, and discusses several other topics.

The second part of the publication concerns applied work on vision. Topics covered include dynamic visual acuity, static visual acuity (in particular, population statistics and changes with age), driver eye height and highway sight distances, field of view and obstructions (e.g., pillars), windshield design, roadway markings, and the relationship between peripheral vision and accidents (especially on expressway ramps).

Several pages are devoted to the design of instrument panel displays. One item examined in detail is the time to look from the road to the instrument panel and back again. These predictions are based on times obtained from aircraft pilots. Unfortunately, the emphasis of the analysis is on complete accommodation, not functional accommodation. Functional accommodation (accommodation sufficient for the task of interest) is most relevant for typical instrument panel display reading tasks. Most of the remaining discussion of instrument design does not refer to any data.

Hence, this document discusses a number of topics related to vision and vehicle design. However, its treatment of the topic of interest, the legibility of displays, is not quantitative. It nonetheless does provide useful background information for designers.

Cornog and Rose (1967)

This 460-page publication is an exhaustive review of the subject of the legibility of text. It includes many items which are obscure and extremely difficult to obtain. Despite its age, it is still an extremely useful document. The bulk of the report consists of reviews of nearly 200 reports, proceedings papers, and journal articles. Each review follows a highly structured format including a complete citation, and sections titled "problem," "procedure," and "result." In many cases, the information in those sections is directly quoted from the source. While it does include some tables of results from the sources, figures are not included. An example of one of the shorter citations follows.

> *3296* *Bridgman-1956*
> *Bridgman, C.S. and Wade, E.A.*
> *Wisconsin U., Madison*
> *OPTIMUM LETTER SIZE FOR A GIVEN DISPLAY AREA*
> *Journal of Applied Psychology, 40:6 (December 1956) 378-380, 5 refs*
>
> *Problem: "Given a certain display space, limited by a high-contrast border, what is the maximally visible size of (an) inscribed ... single line of BLOCK capital letters, (using) a visual acuity criterion of visibility?" (p. 378)*
>
> *Procedure: Forty male and female subjects (psychology students having at least 20/20 binocular visual acuity) viewed five-letter lines of BLOCK, upper case, alphabetic characters projected onto (1) "an aluminized projection screen," (p. 378) and (2) "a flat black mat surface." (p. 378) The experiment was of a minimum recognition size "threshold" design. The two background surfaces provided luminance levels of 8.45 and 0.084 millilamberts, respectively. Ratios of letter size to vertical dimension of the background utilized were 1 to 1, 1 to 1.4, and 1 to 5.5. Using a variable magnification projector, "the relations stated above were maintained as letter size was varied to determine thresholds. ... Projection and observation distance were both 20 feet." (p. 378)*

Result: "Providing a field equal to the stroke width of the letters gave improvement in mean acuity thresholds of nearly 11 (percent) over those obtained with no field, and the wider field (2.25 times the letter size) gave improvements of 18 to 20 (percent). These data are examined in terms of the over-all size of the field required to provide threshold letters, however, it is found that the decrease in letter size is not enough to compensate for the additional space taken up by the field. It is concluded that, when space limitations are a consideration, letters should be made as large as possible up to the point of very nearly filling the available space (margin less than the stroke width of the letters), in order to permit discrimination at a maximum distance." (p. 380)

The report also includes an index, character font samples, a glossary, and a matrix summarizing all of the studies. The matrix appears in Figure 1. As can be seen from the figure, there are several clear trends.

1. Only a minority of the studies varied illumination.
2. Most of the time the display was viewed at console distance (approximately 28 inches).
3. Most studies used numeric characters.
4. Most studies varied font characteristics, in particular height, width, height to width ratio, strokewidth, height to strokewidth ratio, and the choice of font.
5. When layout was varied, several factors associated with it were manipulated.
6. Commonly, the task was recognition of characters.
7. Most articles concerned experiments, as opposed to reviewing the literature of reporting analyses.
8. Accuracy was the most commonly used performance measure.

Readers interested in the topic of legibility are encouraged to read this review.

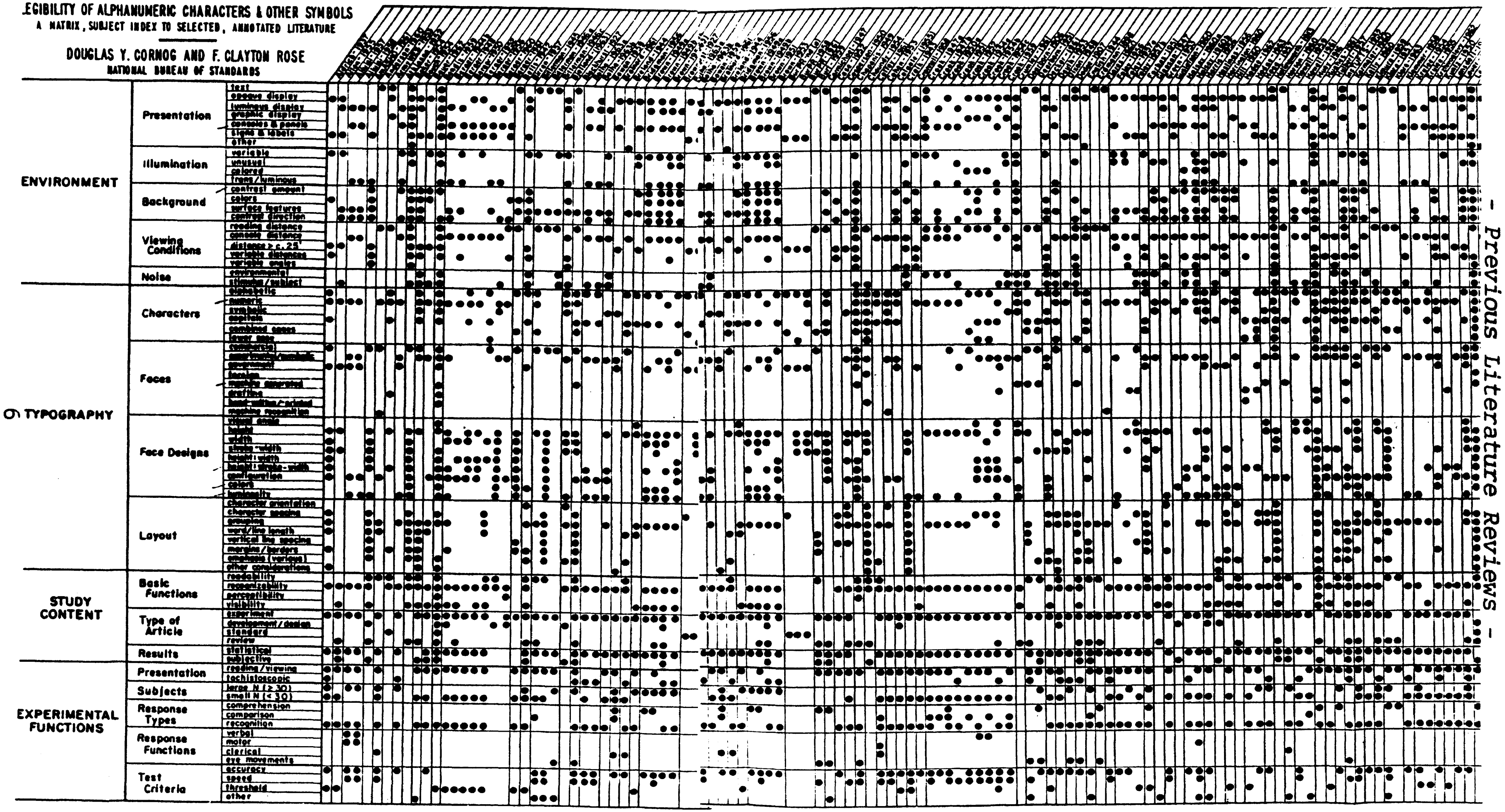
Gallagher, McCunney, and Thornton (1977)

This document describes the results of a literature search on visibility. The purpose of the program was to develop a quantitative index of visibility for NIOSH (National Institute for Occupational Safety and Health). This metric was to be used to evaluate the effectiveness of markings, labels, and signs in warning against occupational hazards. Over 300 manuscripts were reviewed.

LEGIBILITY OF ALPHANUMERIC CHARACTERS & OTHER SYMBOLS

A MATRIX, SUBJECT INDEX TO SELECTED, ANNOTATED LITERATURE

DOUGLAS Y. CORNOG AND F. CLAYTON ROSE
NATIONAL BUREAU OF STANDARDS



ENVIRONMENT	Presentation	text opaque display luminous display graphic display consoles & panels signs & labels other
	Illumination	variable unusual colored trans/luminous
	Background	contrast amount colors surface features contrast direction
	Viewing Conditions	reading distance console distance distance > c. 25' variable distances variable angles
	Noise	environmental stimulus/subject
TYPOGRAPHY	Characters	alphabetic numeric symbolic capitals combined cases lower case
	Faces	commercial experimental/symbolic [illegible] foreign machine generated drafting hand-written/printed machine recognition
	Face Designs	visual angle height width stroke-width height:width height:stroke-width configuration colors luminosity
	Layout	character orientation character spacing grouping word/line length vertical line spacing margins/borders emphasis (various) other considerations
STUDY CONTENT	Basic Functions	readability recognizability perceptibility visibility
	Type of Article	experiment development/design standard review
	Results	statistical subjective
EXPERIMENTAL FUNCTIONS	Presentation	reading/viewing tachistoscopic
	Subjects	large N (≥ 30) small N (< 30)
	Response Types	comprehension comparison recognition
	Response Functions	verbal motor clerical eye movements
	Test Criteria	accuracy speed threshold other

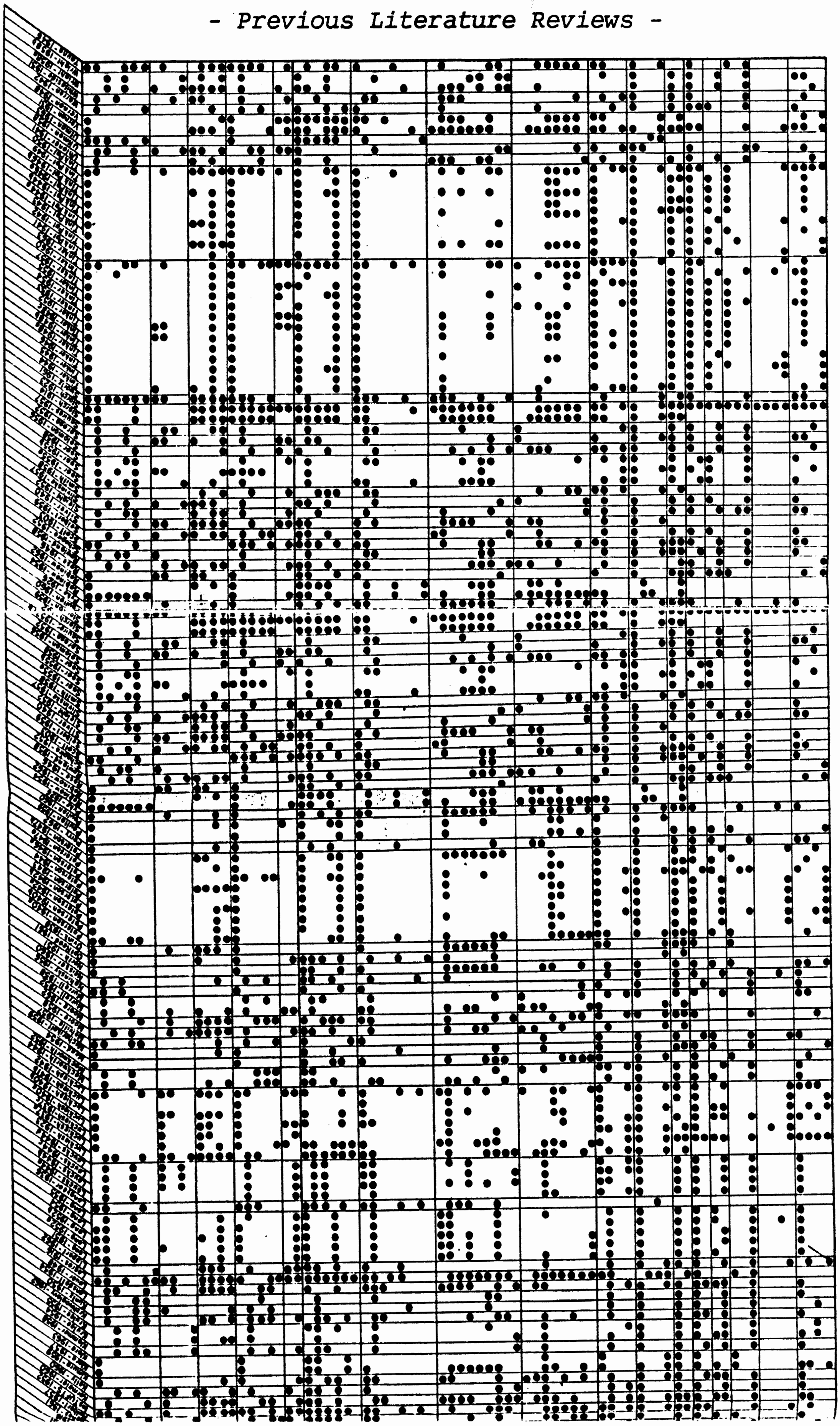


Figure 1. Summary Matrix from Cornog and Rose (1967)

Topics covered include existing standards for visibility, current practice, design factors (camouflage, pattern recognition, color contrast, luminance contrast, area, and reflectivity), and environmental variables (illumination, background, lighting types, distance, and moving objects). Conclusions and research recommendations are also included, neither of which were very specific.

The individual annotations for each reference are quite brief, typically one or two sentences. Figures, tables, and equations are not included. The report therefore serves primarily to flag potential documents of interest rather than providing useful specific technical data. A sample entry follows.

> *106 Copping, B., Alexander, V.D., and Hunter, J.J. Human Factor Assessment of the Legibility of Five Numeric Visual Displays, Applied Ergonomics, 4(3), September, 1973, 144-149.*
>
> *The experiments described were aimed at discovering which of five numerical displays was most suitable to telephonists with a display behind the keyshelf. The results are generally applicable to other office functions. The numerical displays eliminated perceptual confusions of form (0 to o or B and 8) and acoustic confusions in memory of names (B, C, D, and E) with a larger set of symbols.*

Laycock (1985)

This brief Society Information Display paper concerns passive display devices. It assumes the reader has some knowledge both of how passive display devices work and how the human visual system functions. The paper is divided into two parts. The first discusses the physical and optical aspects of displays (dynamic scattering, twisted nematics, etc.) The emphasis of the discussion is on the contrast ratios and device response times associated with various design modifications.

The second part of the paper concerns legibility. Laycock (p. 90) cites the Smith (1979) study as a summary recommendation: "...a 24 minute of arc character height is preferred if speed and accuracy are important, but 12 minutes is sufficient in less critical applications, with 6 minutes representing the lower limit."

Laycock also cites the formulas of Duncan and Konz (1976) concerning the legibility of LCD's under moderate levels of illumination. The derivation of these equations is described in the General Relationships section of this report.

An alternative relationship based on work carried out by Payne (1983) is also included and described later in this report.

Also described is a 1977 study by Buckler, who found that the strokewidth to height ratio could be varied from 1:6 to 1:10 with no significant loss of legibility.

With regard to other design characteristics, Laycock identifies 5x7 dot matrix characters as being "acceptable," 7x9 as being "recommended," and 9x13 as offering "improved performance." These recommendations differ from those of Snyder and Maddox (1978), who suggest the "acceptable" is 7x9, not 5x7. (See the Font section of this report for a discussion of matrix size.) Laycock also notes that slanting characters does not improve legibility.

Finally, Laycock notes small displays (13-16 character window) should scroll pixel by pixel, not letter by letter. Rates of 10-15 characters per second can be read.

Foster (1980)

Foster (1980) has been reviewed in previous reports related to the Chrysler displays project (Adams, Goldstein, Zeltner, Ratanaproeska, and Green, 1988; Zeltner, Ratanproeksa, Goldstein, Adams, and Green, 1988). A review is included here for the sake of completeness.

The aim of Foster (1980) is to provide graphic designers and researchers with a summary of studies published between 1972 and 1978. Nearly 500 citations are included. It does not, however, include any figures or tables. Relevant sections are discussed below.

Sections 2, 3, and 4 deal with processes underlying the comprehension of graphic displays. In particular, section 3 covers experiments on letter and digit identification. A study by Van Nes and Bouma (1977) examined the discriminability of segmented digits and found that digits varied in identifiability (8 being the least discriminable), becoming less identifiable as the number of component line segments increased. This article is described more fully in the Font section of this report. Ellis and Hill (1978) found that more errors were made reading segmented than conventional numerals. However, subjects could be trained to read segmented numerals as proficiently as conventional numerals, but proficiency was lost when retested a month later.

Also of interest are Sections 6 (covering signs, symbols and signing systems) and 10 (covering the presentation of numeric data). Those interested in color coding should look at Subsection 6.10 which discusses Christ's work on the benefits of color coding in visual search tasks.

Previous UMTRI Reviews

As noted in the preface, this report is the third in a series of three written on the subject of legibility. The first (Adams, Goldstein, Zeltner, Ratanaproeksa, and Green, 1988) identifies 121 items in the UMTRI Library pertaining to legibility and includes abstracts for each. The report includes indexes by author, title, UMTRI catalog, which makes the document useful in identifying and locating items of interest.

Based on the information from the first report, 28 of those abstracts were selected for further review in a second report (Zeltner, Ratanaproeksa, Goldstein, Adams, and Green, 1988). Summaries were written for each of the articles which emphasized engineering issues (calculations and applications). Figures and tables were also included. Four key articles were identified for further consideration--Hind, Tritt, and Hoffman (1976) for its coverage of the factors affecting legibility; Van Nes and Bouma (1980) for information on numeral design, Howett (1983) for procedures to calculate letter heights, and Foster (1980) for its review of legibility research. Each of those documents is covered in detail here.

The primary value of the Zeltner et al. (1988) review is that it identifies key relationships and experimental methods. However, because a number of the articles reviewed concern highway signs and license plates, its application to instrument panel displays is not as direct as it could be.

Summary

Those interested in previous reviews of the literature should read the Foster (1980) report first and, time permitting, skim Zeltner et al. (1988) and Cornog and Rose (1967). Readers should be forewarned that because of its structure and length, Cornog and Rose is not light reading. Those with weak backgrounds in human vision may wish to review the Society of Automotive Engineers (1966) document or one of the chapters written by Chapanis in the human factors literature.

EFFECTS OF LUMINANCE CONTRAST AND ILLUMINATION ON LEGIBILITY

The understanding of how luminance contrast (commonly referred to as contrast) and illumination levels affect the recognition of characters is important to the development of a display legibility model. More importantly, this understanding is critical to the design of experiments concerned with legibility. This section reviews several articles which are concerned with both the quantification of some basic legibility factors and two contemporary display image quality metrics (Modulation Transfer Functions and Pixel Error Measurements) which deal with luminance variations in a very nontraditional way. Studies by Blackwell and his colleagues, because they are so substantial, are covered in the section that follows this one.

Cobb and Moss (1928)

This paper examines the four basic factors that affect visual threshold: target size, target luminance, background luminance, and exposure duration. Nine people viewed pairs of vertical bars that were mounted on a disk spinning at high speed. The viewing distance was about 6 m. Periodically, the disk was stopped for either 75, 170, or 300 ms for the participants to view. Background luminances of 1, 20, and 100 mL were examined for several target sizes varying from 0.65 to 16.02 minutes of arc. (The sizes examined depended upon the background luminance.) The contrast was then adjusted to determine the "border of visibility" (Cobb and Moss, 1928, p. 833). Further details are not provided. About 100,000 observations were made.

Shown in Figure 2 is a plot of the results. Notice that both the x and y axes (visual angle in minutes and percent contrast, respectively) are log functions. If linear axes had been used, the departure from linearity shown in the figure would have even been greater.

The data was replotted by Luckiesh and Moss (1937) and appears in Figure 3. The three-dimensional figure clearly shows that exposure duration and target luminance tend to have much less of an effect on performance than does contrast. Notice that it emphasizes the 10 mL background level, which was not a level measured by Cobb and Moss (1928).

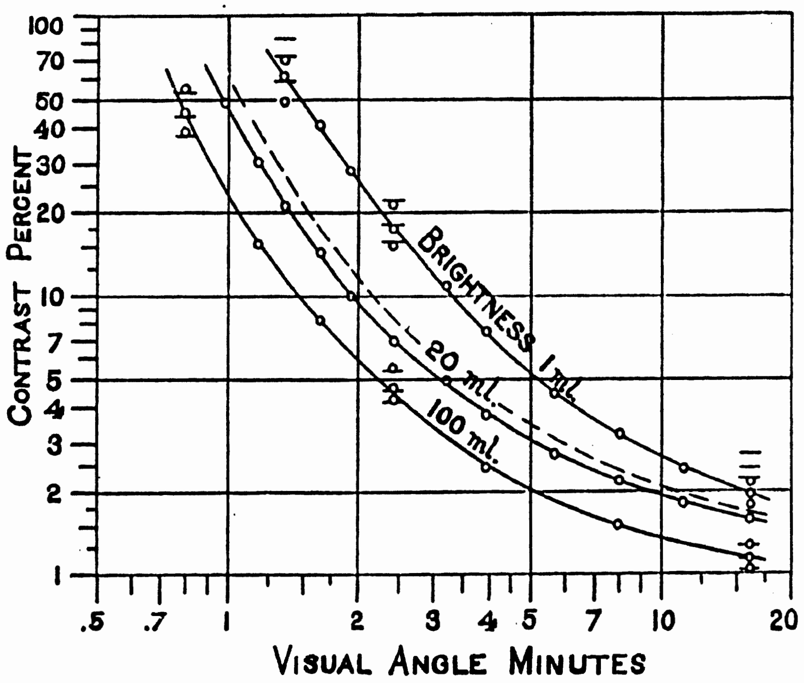


Figure 2. Visual Angle - Contrast Relation from Cobb and Moss (1928)

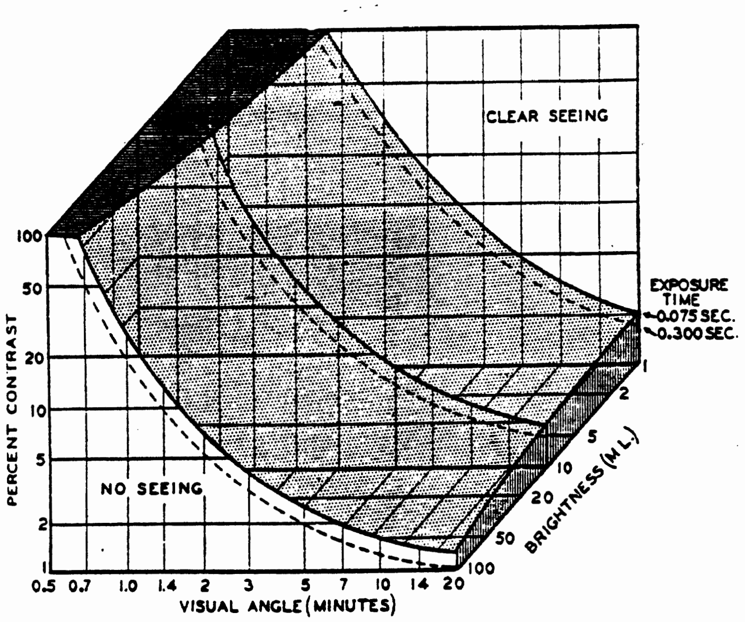


Figure 3. Visual Angle - Contrast Relation from Luckiesh and Moss (1937)

Many human factors specialists have used these data to estimate minimum required character sizes when the lighting parameters (contrast, background luminance) are well specified. In those instances, the visual angle is assumed to correspond to the critical detail (strokewidth) of a character. The character height in those instances is computed by multiplying the strokewidth by the height to strokewidth ratio. To facilitate such computations, the data are often replotted with the x and y axes of the original figure reversed (and the grid lines removed). That, coupled with the difficulty of making log interpolations, makes such calculations error-prone. Figure 4 shows an example of such a plot from Van Cott and Kinkade (1972).

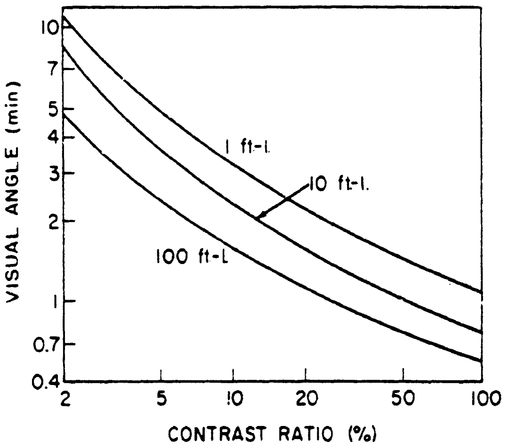


Figure 4. Revised Cobb and Moss Figure from Van Cott and Kinkade (1972)

Moon and Spencer (1944)

The Moon and Spencer document does not report any new experimental work. Rather, it summarizes 80 years of previous research. It identifies the minimum perceptible contrast for large contiguous surfaces with uniform surround luminance and the minimum visual angle required for small objects with high contrast and uniform surround. From this analysis, empirical equations were developed representing average visual thresholds as a function of illumination. Because of its age, this paper has largely been forgotten.

Minimum perceptible contrast is usually defined as the just-noticeable luminance difference between a test object and its background. To determine that threshold, Hecht, Peskin, and Patt, (no date given) had an unreported number of people identify, in an unreported manner, the presence of a target. The

12 degree target was shown on a 40-degree background for 40 milliseconds. A variety of target luminances and contrast levels were examined. Based on this data, Hecht (1935) developed the following equation to predict minimum perceptible contrast:

$$Cmin = (C_{oo}/B) * (A_1 + B^{1/n})^n \qquad (1)$$

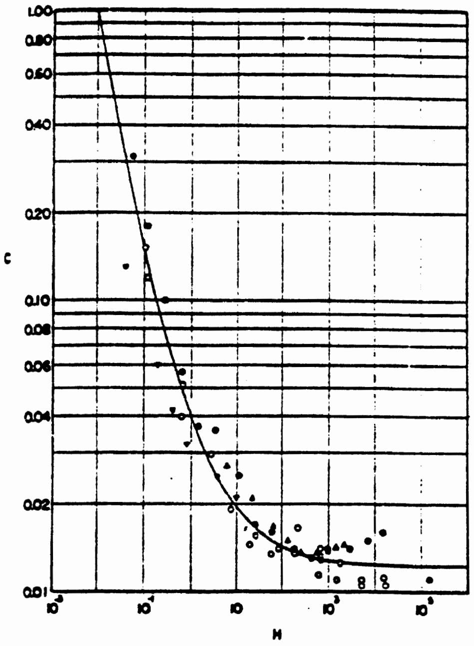
where:

$Cmin$ = Minimum Perceptible Contrast at luminance B
C_{oo} = Minimum Perceptible Contrast for B approaching infinity
A_1, n = constants

As can be seen in Figure 5, minimum perceptible contrast as a function of luminance was plotted based on data from several researchers, and parameters of equation 1 were fitted to these data. The best fit equation was found to be:

$$Cmin = (0.0123/B) * (0.808 + B^{1/2})^2 \qquad (2)$$

It should be noted that this equation applies only to "good" viewing conditions defined as: viewing through a natural pupil on a uniform surround using cone vision; large comparison surfaces (10 - 20 degrees); and exposure times greater than 0.1 second.



Source: Moon and Spencer (1944)

Figure 5. Minimum Perceptible Contrast As a Function of Target Luminance from a Variety of Researchers.

The primary data on visual angle is from Shlaer, Smith, and Chase (1942), as reported in Moon and Spencer (1944). In their experiment, a black Landolt C was viewed against a uniform background through a monocular 2-mm artificial pupil for a variety of ambient conditions. These data were fitted to the form of equation 1, resulting in a relation between character height and luminance:

$$V = (V_{\infty}/B) * (A_2 + B^{1/n})^n \qquad (3)$$

where:

V = Minimum Perceptible Visual Angle (radians)
B = Background Luminance (millilamberts)
V_{∞} = Visual Angle as B approaches infinity
A_2, n = constants

Constants based on numerous data for cone vision using broken circle test objects and large uniform surrounds resulted in the following parameters for equation 3:

$$V = (118.3X10^{-6}/B) * (0.412 + B^{1/3})^3 \qquad (4)$$

where:

V = Minimum Perceptible Visual Angle (radians)
B = Luminance of the Background and Surround (millilamberts)

Nothing is mentioned by the authors about target luminance in association with this equation. Equation 4 can be used to determine either the minimum size object that can be seen at a given distance, the maximum distance at which a given object can be seen, or the luminance required for a given visual angle.

For single-bar test targets of considerable length, equation 3 should be modified as follows:

$$V = (1.950*10^{-6}/B) * (1.433 + B^{1/2})^2 \qquad (5)$$

It should be noted that equations 3-5 apply to observers with excellent vision under very favorable viewing conditions, although these conditions were not specified. Further, V in these equations was defined so that the visual angle subtended by the target will be such that the percent of responses correct is 56.25%. Equations to predict V for different values of percent correct are as follows:

$$V = (V_{oo}/B)*((y-1/8)/(1-y))^{1/k} * (A_2 + B^{1/n})^n \qquad (6)$$

where:

V = Visual Angle (radians) corresponding to y
y = Fraction of Total Answers that are correct
k, n, A_2 = constants.

Using constants as before for Landolt-type test objects, the above equation becomes:

$$V = (118.3*10^{-6}/B)*((y-1/8)/(1-y))^{0.01} * (0.412 + B^{1/3})^3 \qquad (7)$$

Equation 7 is the most useful result from the Moon and Spencer article. It is a general equation which can be used to determine visual angle, required viewing distance, or required letter height for visibility for a given percent correct criterion. Further, it is applicable to Landolt-type test objects, which approximate legibility requirements of alphanumeric characters very well.

One important result based on equations 6 and 7 is that if a certain object is visible only half the time when flashed onto a field of a given luminance, doubling the size of the object will make the object visible all of the time. This result has been found by other researchers (Blackwell, 1946), and is useful in the design of legibility experiments. Since it is easier and more time efficient to collect data on 50% threshold letter size than it is for error-free letter size, this rule of thumb allows for simpler data collection procedures to be used.

Snyder and Maddox (1978)

This report describes an extensive series of experiments which assessed legibility of dot matrix displays and the relationship between legibility and character appearance variables. This report has been included in this section because it examines performance consistency as lighting variables are manipulated.

The first experiment examined the relationship between performance measures and legibility. (See also Snyder and Taylor, 1979.) Six people were shown 7x9 dot matrix characters on a CRT. The performance measures included: 1) accuracy of response under normal viewing conditions; 2) response time to identify single characters; 3) tachistoscopic recognition; and 4) threshold visibility. For the accuracy and response time measures, there were three luminance levels (8, 27, 80 cd/m^2), four character sizes (2.64, 3.05, 4.79, and 5.44 mm), and seven viewing distances (0.61 to 3.35 m). For the tachistoscopic recognition data, four character sizes were shown at three luminance levels and three exposure durations (17, 33, and 50 msec). The 50% and 85% visibility thresholds were collected for all 12 character size/luminance combinations.

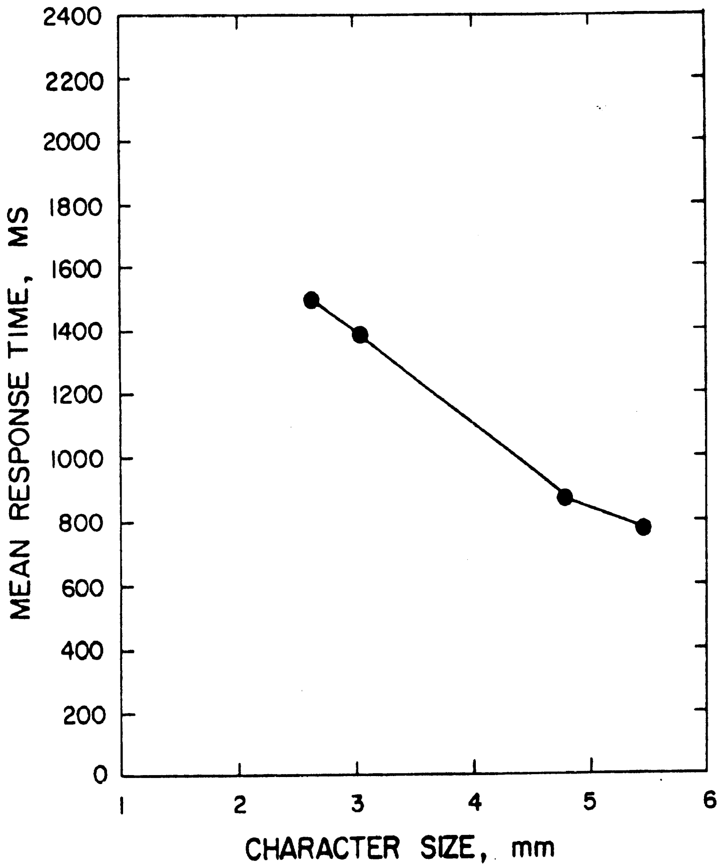
Table 1 shows the correlations between performance measures based upon the 12 character height/luminance means. The two highest correlations were between accuracy and threshold visibility, and response time and threshold visibility. Correlations of the tachistoscopic recognition measure with other measures improved as the duration was decreased. Snyder and Maddox use these results, as well as other evidence, to argue for using accuracy as a performance measure in subsequent studies. They also support the use of classic tachistoscope measures reported earlier for no-search tasks.

Table 1. Correlations Between Response Measures in Snyder and Maddox (1978).

Measure	Accuracy	Response Time	Tachistoscopic Recognition 17 ms	Tachistoscopic Recognition 33 ms	Tachistoscopic Recognition 50 ms
Response Time	-.97				
Tachistoscopic Recog. (17 ms)	.78	-.72			
Tachistoscopic Recog. (33 ms)	.83	-.64	.62		
Tachistoscopic Recog. (50 ms)	.60	-.60	.38	.53	
85% Threshold Visibility	.94	-.96	.69	.78	.48

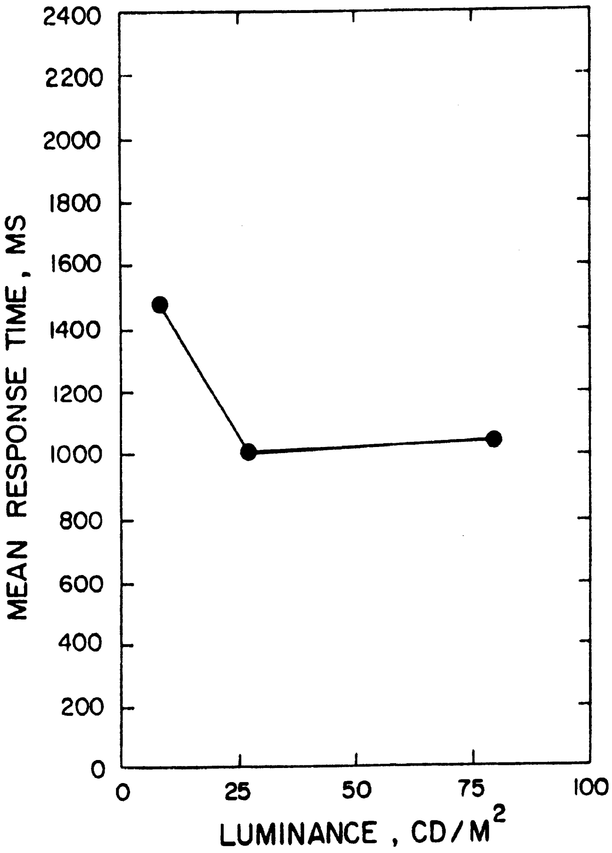
The Snyder and Maddox report includes numerous figures showing the relationship between the various independent and dependent legibility measures described above. Two figures of particular interest are the relationships between response time and character height (Figure 6), and response time and character luminance (Figure 7). The relationship between size and response time is inverse and linear except for the largest character size. The relationship between response time and luminance is very nonlinear.

Snyder and Maddox conducted a third series of experiments to determine if the Modulation Transfer Function (MTF) of a display could be used to assess image quality. (MTF and other related measures are described at the end of this section.) Four-letter anagrams (e.g., AESY) or words (e.g., EASY) were presented for 17 ms each on a CRT to six people. The viewers' task was to recall the characters shown. The four character sizes (2.64, 3.05, 4.79, and 5.44 mm) were each shown at three illumination levels (8, 24, 66 cd/m2, (2.3, 7, 19 fL)). The viewing distance was 0.61 m.



Source:
Snyder and
Maddox (1978)

Figure 6. Effect of Character Height on Response Time.



Source:
Snyder and
Maddox (1978)

Figure 7. Effect of Character Luminance on Response Time.

The word data were messy with few significant interactions. On the other hand, for the anagram data there were significant differences due to character size, luminance, and the size-luminance interaction. In general, the differences between the word and anagram scores increased as character sizes decreased. Because of this, it can be concluded that context cues are far more important for poor viewing conditions.

The most important finding in this experiment was the good correlation ($r > 0.8$) between the weighted log Modulation Transfer Function Area (MTFA) data and the log percent correct for anagram recognition. (See Snyder and Maddox, 1978 for a description of the calculation procedure. MTFA and other measures are described later in this section.) This finding points to the possibility of the evaluation of image quality from physical measurements of displays alone. These and other related measures are described in greater detail in the next section.

In a subsequent experiment, simulated characters from three electronic displays were shown on a high-resolution CRT. These displays all had either triangular or rectangular matrix elements. The first phase involved 5x7 dot matrix characters and 72 participants. In tasks similar to those described in the first experiment, people searched text for inappropriate words, searched for targets in structured arrays, and searched for characters in unstructured arrays. Participants pressed a button when they found the inappropriate word or target character and the time was recorded.

The second experiment involved 5x7, 7x9, and 9x11 dot matrix characters and 40 participants. Performance measures included reading speed and times to locate targets in a structured (menu) search. The tasks performed were identical to those in the previous experiment. From a stepwise regression analysis, Snyder and Maddox developed the following models.

Tinker Reading Task

$$\text{Adjusted Reading Time (s)} = 1.43 + 0.23(\text{VSQR}) + 3.64(\text{HMTFA}) + 0.221(\text{VMTFA}) - 4.825(\text{HMLOG})$$

Menu Search Task

$$\text{Search Time (s)} = 0.78 + 0.024(\text{VSQR}) + 2.72(\text{HLOG}) + 0.193(\text{VMTFA})$$

Random Search Task

$$\text{Search Time (s)} = -48.50 - 138.49(\text{HFLOG}) + 192.89(\text{VFLOG}) - 0.642(\text{HMTFA}) - 0.734(\text{HSQR}) + 0.982(\text{VSQR}) - 0.043(\text{HDIV})$$

where:

VSQR, HSQR = Vertical, Horizontal Square of Fundamental Spatial Frequency Minus 14.0
VMTFA, HMTFA = Vertical, Horizontal Pseudo-modulation Transfer Function Area
VMLOG, HMLOG = Log10 of VMTFA and HMTFA
VDIV, HDIV = Fundamental Spatial Frequency Divided by Modulation
VLOG, HLOG = Log10 of VDIV and HDIV
HFLOG, VFLOG = Log10 of Fundamental Spatial Frequency

These equations accounted for 50-60% of the variance of the search/reading times (as appropriate).

With regard to the effects of font on performance, there were no differences between the 7x9 and 9x11 fonts on reading performance, but both were surprisingly inferior to the 5x7 matrix. For the menu search task, increases in matrix sizes led to consistent improvements in performance.

A subsequent experiment concerned validation of the models developed. Three working plasma panels (rather than simulations of them examined previously) were used in menu search and reading tasks identical to those in the previous experiment. The performance measures were also the same as before. Seventy-two people participated. The ordinal relationships between the predicted and actual performance scores were the same, although the actual values deviate noticeably from the predicted values.

Snyder (1985)

This study presents several measures of display image quality proposed by various researchers. They are quite different from the classic measures of legibility described earlier and are likely to be unfamiliar to readers. Their primary use has been for the evaluation of display systems hardware rather than particular images (e.g., character fonts). These image quality metrics are divided into two classes: Modulation Transfer Function measurements and Pixel Error measurements.

The Modulation Transfer Function (MTF) plots modulation transfer factor versus spatial frequency in cycles per degree of visual angular subtense. The modulation transfer factor is defined as: Modulation-out/Modulation-in. Modulation, which is basically an expression of luminance contrast, is defined as:

$$\text{Modulation} = (L_{max} - L_{min})/(L_{max} + L_{min})$$

where:

L_{max} = Luminance of the Lighter Sine-wave Grating Half Cycle
L_{min} = Luminance of the Darker Sine-wave Grating Half Cycle.

The sine-wave grating to determine Lmax and Lmin is used by first choosing a spatial frequency (cycles per unit display distance or cycles per visual degree of angular subtense) and then adjusting the luminance contrast (modulation) of the grating to a threshold criterion. In terms of image quality, the MTF relates the modulation input of a device to the modulation output of that device for a sine-wave input of a given spatial frequency and modulation.

Several specific MTF based measures of image quality are briefly enumerated as follows. For a more indepth discussion of these specific MTF measures, the reader should consult the actual research paper cited in the reference section of this report.

> Equivalent Passband (Schade, 1953): Based on the variance of a distribution, this method measures the blur or sharpness of an image.
>
> Strehl Intensity Ratio (Linfoot, 1960): A modification of the equivalent passband method using different weighting schemes for the MTF.
>
> Modulation Transfer Function Area (MTFA): This measure, proposed by Charman and Olin (1965), finds the area between the MTF function and the Contrast Threshold Function (CTF), as seen in Figure 8. (The CTF is found when modulation threshold contrast is plotted as a function of spatial frequency.) This measure shows, for a given spatial frequency, the amount that a system output signal is above the threshold contrast of the human visual system. A different interpretation of this measure is that it shows the difference between overall signal intensity (MTF) and signal noise (CTF).

(Figure 4.4 from Snyder)

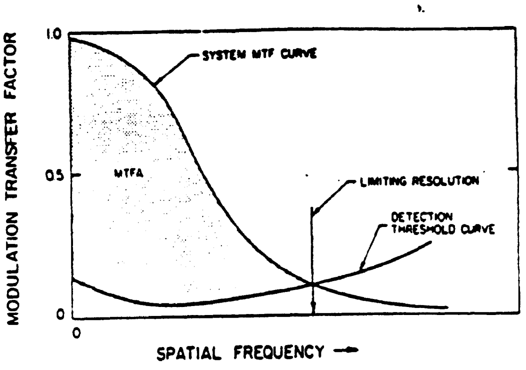


Figure 8. Example of the MTFA Relationship to MTF and CTF.

> Gray Shade Frequency Product: Beamon and Snyder (1975) found that as modulation is increased away from the CTF curve on the MTFA, the rate of increased visibility decreases rapidly. Therefore, the area directly above the CTF curve on the MTFA is much more important to observer visibility than the area far above the CTF curve. To account for this finding, Task and Verona (1976) proposed a transform to the MTFA to weight the area near the CTF curve more heavily than the area closer to the MTF curve. The transform used assumes that the visual system acts as a logarithmic amplifier which sees modulation proportional to the logarithm of the modulation.
>
> Integrated Contrast Sensitivity (Van Meeteren, 1973): This measure provides a different weighting scheme for the system MTF so that it is more sensitive than the MTFA to small changes in the MTF or CTF curve, and consequently, more sensitive to small changes in image quality.
>
> Discriminable Difference Diagrams (Carlson and Cohen, 1978): This method predicts the increase in modulation necessary to achieve a just noticeable modulation difference as a function of spatial frequency.
>
> Displayed Signal-to-Noise Ratio (Rosell, 1971): This method, used for analyzing television systems, produces signal-to-noise ratios which are related to probability of detection. These ratios can in turn be used as standards for developing system hardware requirements.
>
> Visual Efficiency (Overington, 1976, 1982): This is a mathematical model to predict human visual performance, and is used to determine whether the perception of image detail is limited by the physical capabilities of the eye or by the system sharpness of the image.
>
> Information Content (Shannon and Weaver, 1949): This method employs information theory to derive an expression relating information content of pictorial displays to spatial frequency.

The second measure of image quality involves pixel error measures. These measures compare, pixel by pixel, the degradation in intensity level between the original image and the image which is being displayed by some system. These degradations are averaged in some way so that they can be used as an overall index of quality. The primary pixel error measure is the normalized Mean Square Error. This measure is the unweighted sum of the normalized squared deviations between the input image and the system output image. Other pixel error measures, which are modifications of the Mean Square Error measure, are summarized by Pratt (1978) and include Point Squared Error, Perceptual Mean Square Error, Image Fidelity, Structural Content, and Correlation Quality. Interested readers should consult the Snyder (1985) article for further explanation and equations not included in this review.

Summary

The Cobb and Moss study identifies the legibility threshold for simple targets as a function of luminance contrast, background luminance, and exposure duration. This work has been summarized as a series of curves ("Cobb and Moss Curves") which still see widespread contemporary use. Coupled with knowledge of a character set's strokewidth and height to strokewidth ratio, this information can be used to compute the associated legibility distance or required letter height for that target.

The Moon and Spencer studies were concerned with the just noticeable difference between contrast of the target and background. They also report legibility thresholds for targets in terms of visual angle. Just as with the Cobb and Moss (1928) results, these results can be used to determine absolute lower bounds for the contrast needed for legibility. A comparison of letter heights predicted by this method and several others appear in the Application Independent Studies of Legibility section of this report.

Snyder and Maddox provide a number of interesting general relationships between legibility factors. Most important of these include the inverse linear relationship between response time and character height, the high correlation between accuracy and threshold visibility, and response time and threshold visibility, and the non-linear relationship between response time and luminance. This latter relationship has been found to be linear when response time is plotted against log luminance (Smyth, 1947; Forbes, Saari, Greenwood, Goldblatt, and Hill, 1976).

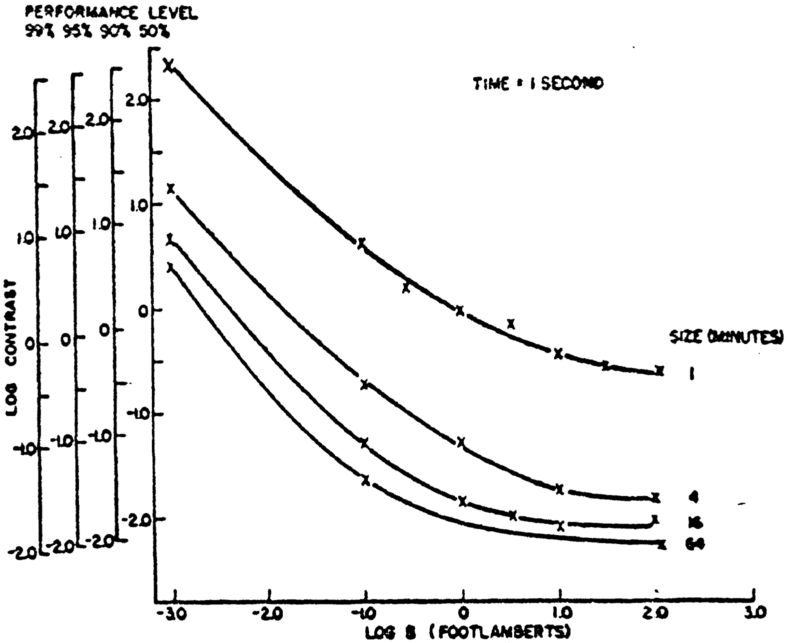
In terms of the image quality metrics described by Snyder, the Modulation Transfer Function Area is seeing increasing use because it allows for the common evaluation of the capabilities of the user and the display in similar units before the complete system is built. Further, Snyder and Maddox provide data which suggest that image quality can be evaluated purely on physical measurements of the display in question by using the relationship between the weighted log MTFA and reading percent correct data. However, the utility of this system is currently limited due to the complex mathematical relationships involved, and because few people can interpret the numbers this method produces, and understand spatial frequency and Fourier transforms.

BLACKWELL'S RESEARCH

This section describes the work conducted by Blackwell over many years. His research examines the relationship between illumination levels, contrast, target luminance, viewing time, and a host of other factors on the detection of simple visual targets such as discs or Landolt rings. His work is detailed, carefully done, and comprehensive. The discussion here only touches upon what he has done.

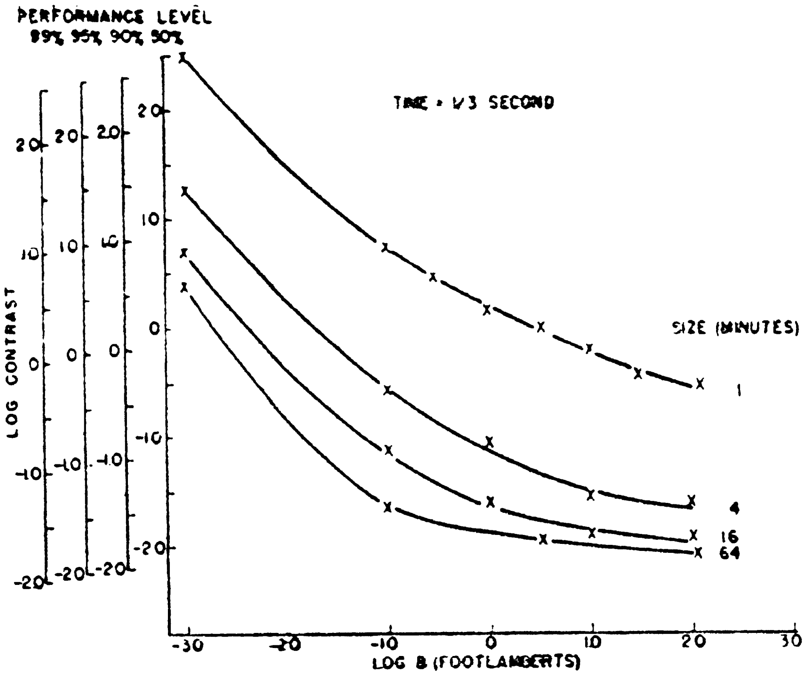
Blackwell (1952)

The paper describes an experiment in which two people were shown discs varying in size from 1 to 64 minutes of arc. Their task was to determine in which of four brief time periods the luminance gradient was shown. Discs were always brighter than the background. The background luminance varied from .001 to 100 ft-L, and exposure duration from .001 to 1 second. For each participant, 250 responses were obtained for each combination of conditions resulting in a total of 81,000 observations. Shown in Figures 9 through 12 are the results for exposure durations of 1 second, 1/3, 1/10, and 1/30 second respectively.



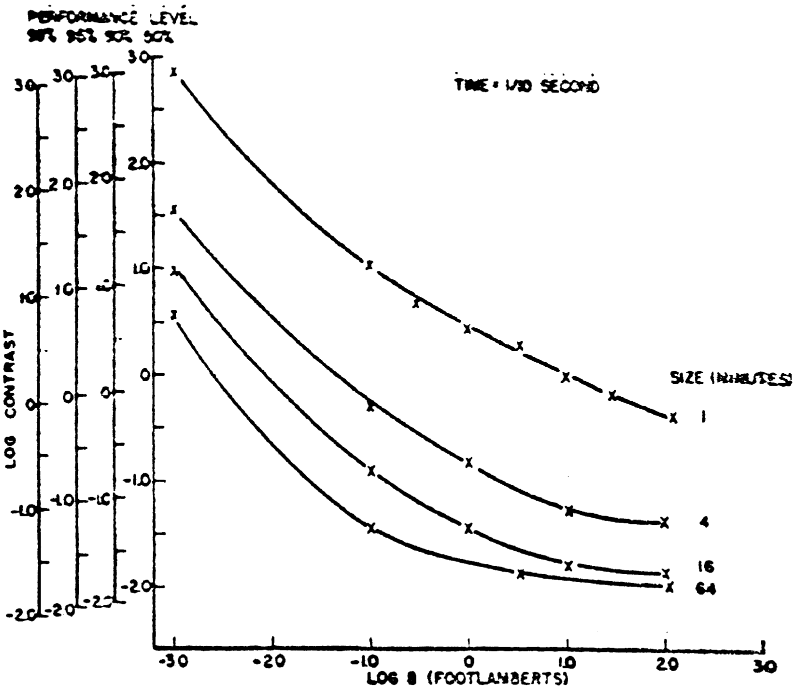
Source: Blackwell (1952)

Figure 9. Log Brightness Versus Log Contrast for 1 Second Exposures.



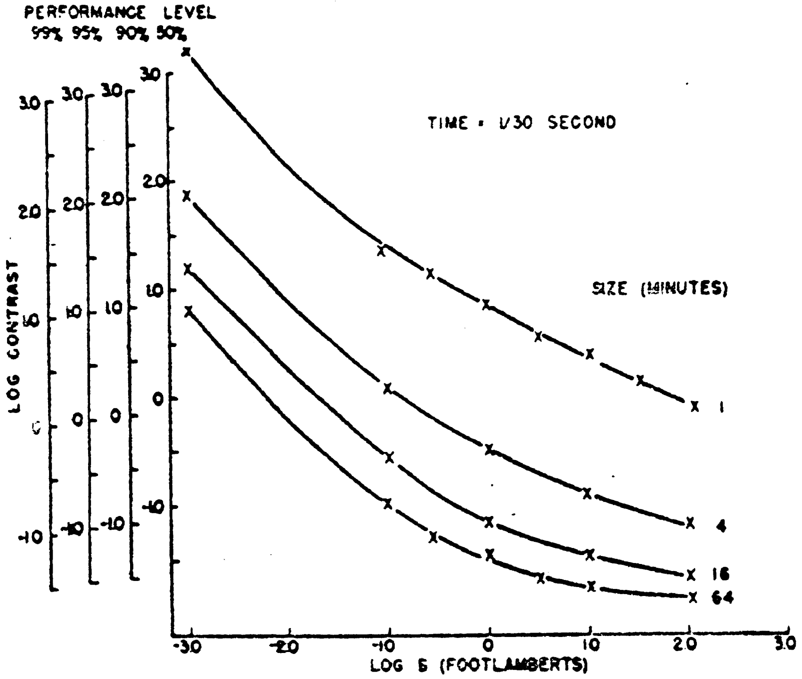
Source: Blackwell (1952)

Figure 10. Log Brightness Versus Log Contrast for 1/3 Second Exposures.



Source: Blackwell (1952)

Figure 11. Log Brightness Versus Log Contrast for 1/10 Second Exposures.



Source: Blackwell (1952)

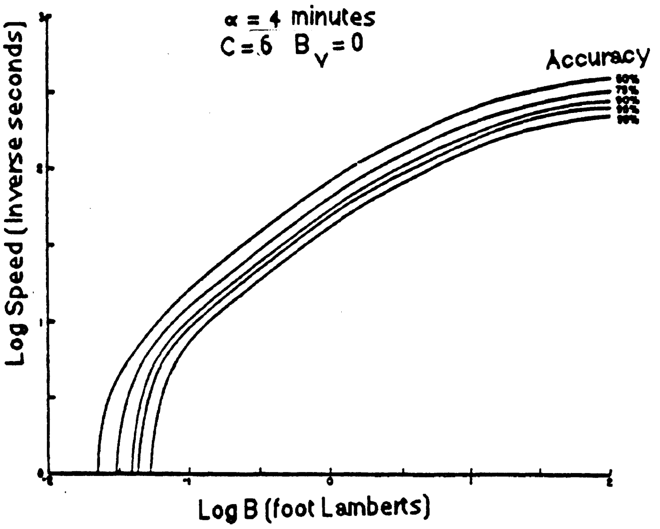
Figure 12. Log Brightness Versus Log Contrast for 1/30 Second Exposures.

Notice that the relationship between log background brightness and log contrast is similar for all exposure durations. In general, the relationship is fairly linear, with the departure from linearity increasing for target sizes above 4 minutes of arc and background lighting levels above 1 foot-Lambert.

Blackwell (1955)

This report discusses a method to determine interior illumination requirements to meet desired speed and accuracy criteria. The method is based on performance data collected during a previous experiment (Blackwell, 1952) for which additional analysis was completed. The paper examines two cases, uniform and nonuniform visual fields. They differ in that the nonuniform case includes veiling glare.

The method assumes that performance can be estimated for conditions not examined by using effective contrast (based on accuracy and time constraints) to extend the data from the test conditions. Figure 13 is an example of the results that emerge from this approach. In these figures, speed is defined as the inverse of the exposure duration for a single presentation during the data collection procedure.



Source: Blackwell, 1955

Figure 13. Interpolated Visual Performance Curves for 4 Minutes of Arc Target, Contrast=.6

For the more common non-uniform luminance case, readers should see the Blackwell (1955) paper.

Farber (1988)

This report describes improvements to a computerized system to predict seeing distances to objects illuminated by vehicle headlamps. The computer program, called DETECT, uses relationships based on Blackwell's work (described previously). The current version computes the detection threshold as follows.

They are based upon the equations given in CIE Publication 19/2.1.

Contrast Threshold = Cth = $Cx * (0.923/n) * [(S/t*Lb)^{.4}+1]^{2.5})$

Where:
Cth = Threshold Contrast
d = Target Diameter (minutes of arc)
Cx = Target Size Factor
if d<=10, $Cx = 3*(.37)^{\log_2(d)}$
if d>10, Cx = .106 - .0006d
Lb = Background Luminance (cd/m^2)
n = $((S/100t)^{.4} + 1)^{2.5}$
S = 0.5900 - 0.6235log d - s
(effect of Age on slope of RCS function of luminance)

s = Adjustment Parameter
if Age 20-44, s=0
if Age 44-64, s=.00406 (A-44)
if Age 64-80, s=.0812 + .00667 (A-64)

t = Relative Equivalent Ocular Transmittance
(loss due to age)
if Age 20-30, log t=0
if Age 30-44, log t=.01053 (A-30)
if Age 44-64, log t=.1474 -.0134 (A - 44)
if Age 64-80, log t=.4154 - .0175 (A - 64)

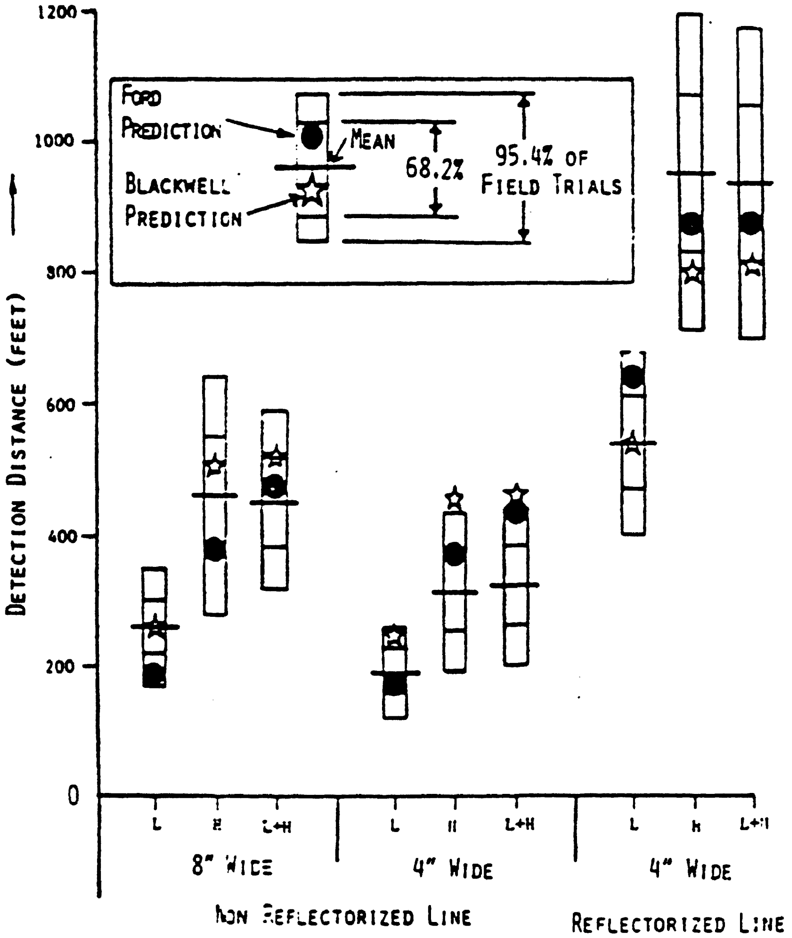
The critical group of terms in the model is the Relative Contrast Threshold (RCS), an expression used to adjust to fit a basic detection function to a wide variety of circumstances. It is expressed as:

RCS = n * [(S/tL)^.4 + 1]^-2.5

Finally, to adjust for the change in contrast threshold with age, Farber describes a multiplier used by Blackwell as follows:

if Age 20-42, m1 = 1.000 + .00795 (A - 20)
Age 42-64, m1 = 1.175 + .0289 (A - 42)
Age 64-80, m1 = 1.811 + .1873 (A -64)

One of the advantages of the Farber paper is that it shows how well the Blackwell model can be used to predict visual performance for a complex task, such as detecting road edge markings. Shown in Figure 14 is a comparison of the model data with on-the-road performance.



Source: Farber (1988)

Figure 14. Comparison of Predictions from Model Based on Blackwell's Data with Field Observations.

Admittedly, the description given here of Blackwell's current model is brief. A more complete description appears in Commission Internationale de l'Eclairage (1972). Those interested in additional details should see Blackwell and Blackwell, 1980a, 1980b; Blackwell, 1981; and Blackwell, 1982.

Summary

Blackwell's work rather completely defines contrast threshold (detection) as a function of a variety of factors. It is extremely complex, quite difficult to understand, and unfortunately has not found its way into contemporary human factors textbooks. A second difficulty in applying it specifically to instrument panel legibility questions is that it ignores visual search of complex fields (looking for the text), which is an important factor. More recent versions of the model, however, do include adjustments for target eccentricity. When legibility comparisons are made, Blackwell's work and his predictions for contrast thresholds should be considered. Readers interested in other perspectives of this work should see Clear and Berman (1983) and Rea (1986).

EFFECT OF FONT ON LEGIBILITY

There have been an extremely large number of studies concerning the effect of font on reading performance. For example, Cornog and Rose (1967) lists several hundred. While the effects of font on performance are less profound than those due to contrast or illumination level (see Hind et al. 1976), fonts are readily modified and hence a subject of intense interest. This section concerns both dot matrix and stroke matrix (segmented) characters. The latter format is given more emphasis because the primary application of the research in this project is toward speedometer design, which presently employs stroke matrix digits. This emphasis is in contrast to the literature, which has concerned itself primarily with dot matrix characters.

Plauth (1970)

Most aircraft displays use characters that conform to MIL-STD-803A-1 (originally referred to as AMEL numerals, now known as the NAMEL or Naval Air Materiel Equipment Laboratory numerals). Plauth examined that font and two others--a standard set of 7-segment numerals and a set slanted 15 degrees to the right. (See Figure 15.) Three sets of 36 slides of five-digit numbers were projected on a screen mounted in an aircraft instrument panel. All digits were 3/8 inch high and viewed at 30 inches. The height to strokewidth ratio was 8:1 for the NAMEL numerals, and about 16:1 for the segmented numerals. The numerals were white characters on a black background.

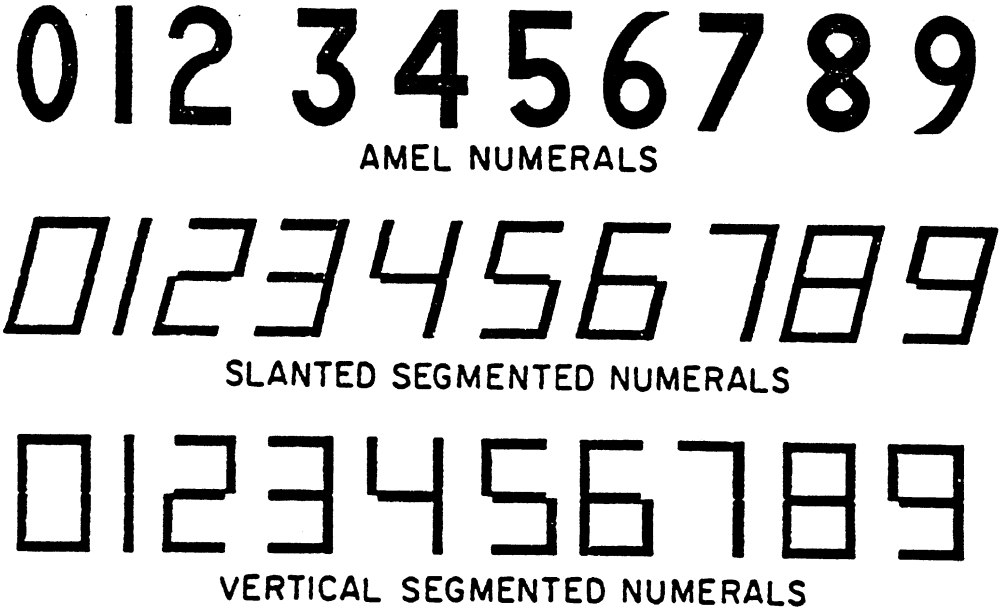


Figure 15. Numerals Examined by Plauth (1970)

The 30 college students who participated were divided into three groups, viewing all fonts for either 1/50, 1/10 or 1/2 second each. Their task was to write down the number displayed on the screen.

Shown in Table 2 are the error data. Reading errors were significantly fewer for both the 1/2 second exposure duration and the conventional (NAMEL) numerals. The poorer performance of segmented numerals is consistent with other findings in the literature. (See Gibney, 1967.) From this, the authors conclude that "segmented numerals, as presently designed, should not be used in applications where accuracy is critical and exposure time is severely limited" (Plauth, 1970 p. 496). One implication of this statement is that fonts other than 7-segment should be examined for automotive speedometers. This does not mean, however, that numeric speedometers should not be used in cars. In fact, as shown in Green (1988), drivers make fewer errors and spend less time reading numeric speedometers than conventional moving pointer displays.

Table 2. Error Data for Plauth (1970).

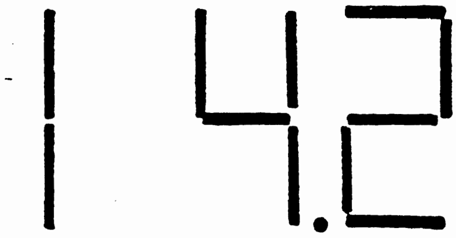
Font	Exposure Duration (sec) 1/50	1/10	1/2	Total Errors
NAMEL	60	90	29	187
Slanted	174	181	36	391
Vertical	158	177	53	388
Total	400	448	118	966

Automobile instrument panel designers often want to use nonvertical segmented numerals to enhance cluster appearance. Plauth found that the difference between the slanted and vertical segmented displays was not significant (3 errors). Thus, using slanted numerals should be acceptable.

Finally, Plauth notes that several people in the experiment complained about the "gap" problem associated with the numeral "1" (See Figure 16). While reading error data for this problem are not provided, display designers should keep in mind that users find its appearance unappealing.

Smith (1978)

When designing a set of characters, two issues must be addressed. First, the design of a character must match the reader's expectation for how that character should appear. Second, the characters should be designed so that they can be easily discriminated from one another. One character set, designed specifically to maximize discriminability while leaving reader expectations as a secondary issue, is the Lansdell set shown in Figure 17.



Source: Plauth (1970)

Figure 16. Illustration of the "Gap" Problem for Segmented Displays.

Source: Smith (1978)

Figure 17. Lansdell Number Set.

While the initial work on the Lansdell set was quite promising in terms of error-free readability, Smith (1978) shows trends to the contrary. One hundred twenty people were shown pages on which 20 rows of 5-digit numbers appeared. Numbers were in either the Lansdell, Mackworth, or Elite font. All numerals were 11 mm high. Figure 18 shows the Mackworth and Elite fonts.

In the first task, participants placed a check mark beside each group containing a target digit. In the second task, participants identified the five-digit groups whose sum was more than 23. In general, the Lansdell numerals were scanned 15-18% more slowly and added 34% more slowly than the other fonts. These differences were both statistically significant. It can therefore be concluded that there are bounds beyond which enhancing numeral discriminability through changes in font can be counterproductive.

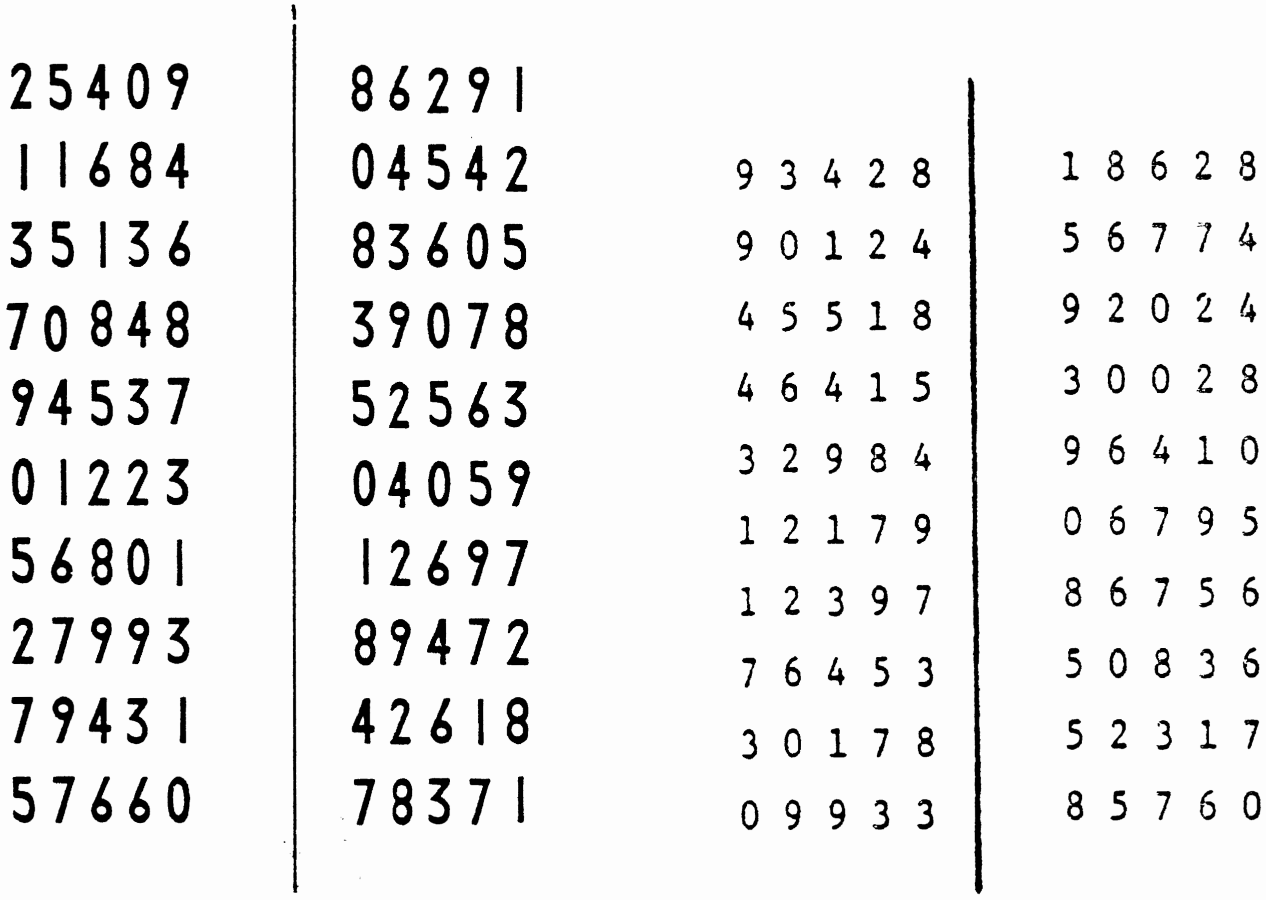
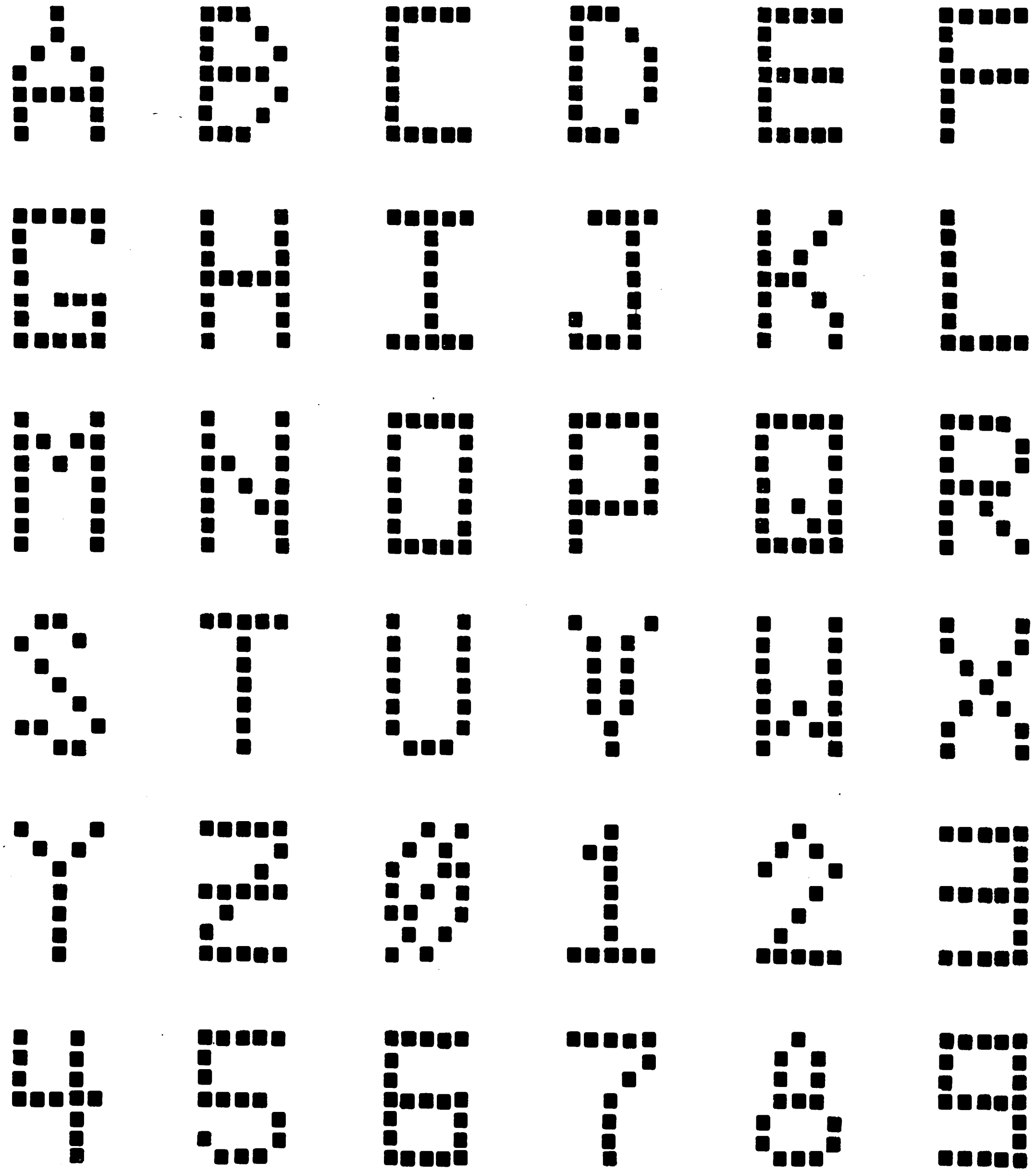


Figure 18. Mackworth and Elite Fonts.

Snyder and Maddox (1978)

One experiment performed by the authors concerned finding optimal dot size-shape-spacing combinations for 5x7 dot matrix characters as a function of ambient illumination. There were three dot (element) shapes (square, horizontally elongated, vertically elongated), three element sizes (0.76, 1.14, 1.52 mm), three between element spacing/element size ratios (0.5, 1, 1.5) and two ambient illumination levels (5.4, 700 lux (0.5, 65 fc)). The font examined is shown in Figure 19.



Source: Snyder and Maddox (1978)

Figure 19. Font Examined in Dot Parameter Experiment.

A total of 108 college students participated with each one seeing only a subset of the test conditions. Each participant carried out several tasks. In a reading task, passages of text were shown. When a word that did not make sense appeared (e.g., "music" in "I wish he had set off his firecrackers at home, for it is too much music for me."), the participant pressed a key and said the word. The performance measure was the number of passages read per five minutes.

In a menu search task, 8x3 matrices of five characters were shown. When the participants found a five-character string that matched the target, they pressed a button. The performance measure was time per trial.

In a random search task, 71 characters were randomly distributed across the screen. Every character appeared twice except for the target. When the target was found, the participant pressed a button and identified the character's location. Here too, the performance measure was time per trial.

With regard to the various test procedures, the reading task was the most sensitive to dot element design. Further, based on the data, Snyder and Maddox recommend that the dots should be square, not rectangular, and that performance was best when the between dot spacing was minimized. The character size preferred depended upon the task. For the reading task, performance was best with the smaller characters. For the search task, larger characters optimized performance. The illumination level produced an overall effect on legibility, but did not affect dot design parameters for the two illumination levels examined.

Subsequent research by Snyder and Maddox focused on font optimization. The first experiment (see also Maddox, Burnette, and Gutmann, 1977) concerned 5x7 dot matrix characters, quite common at the time. Three fonts were examined, as shown in Figure 20. The Lincoln/Mitre font was used because it was consistently reported in the literature as being the most legible font. The Maximum Dot font was constructed by utilizing as many dots as possible in a 5x7 field, which results in a boxy appearance. In contrast, the Maximum Angle font was constructed using as few dots as possible, while emphasizing the use of angles in its construction.

Twenty college students were shown characters one at a time on a CRT. The exposure duration was 40 msec. Their task was to press a key corresponding to the character shown.

Significantly fewer errors were made in responding to the Maximum Dot font (658) than the Lincoln/Mitre (789) or Maximum Angle fonts (764). The difference between the Lincoln/Mitre and Maximum Angle fonts was not significant. The report also presents a detailed analysis of the character confusions.

In a follow-up experiment, error data were obtained for four fonts (Lincoln/Mitre, Maximum Angle, Maximum Dot, Huddleston) for three matrix sizes (5x7, 7x9, 9x11). The procedure was identical to the previous experiment. While this experiment may seem like a needless replication, it was carried out to investigate if an equipment flaw had biased the results from the previous experiment. Several of the character sets examined are shown in Figure 21.

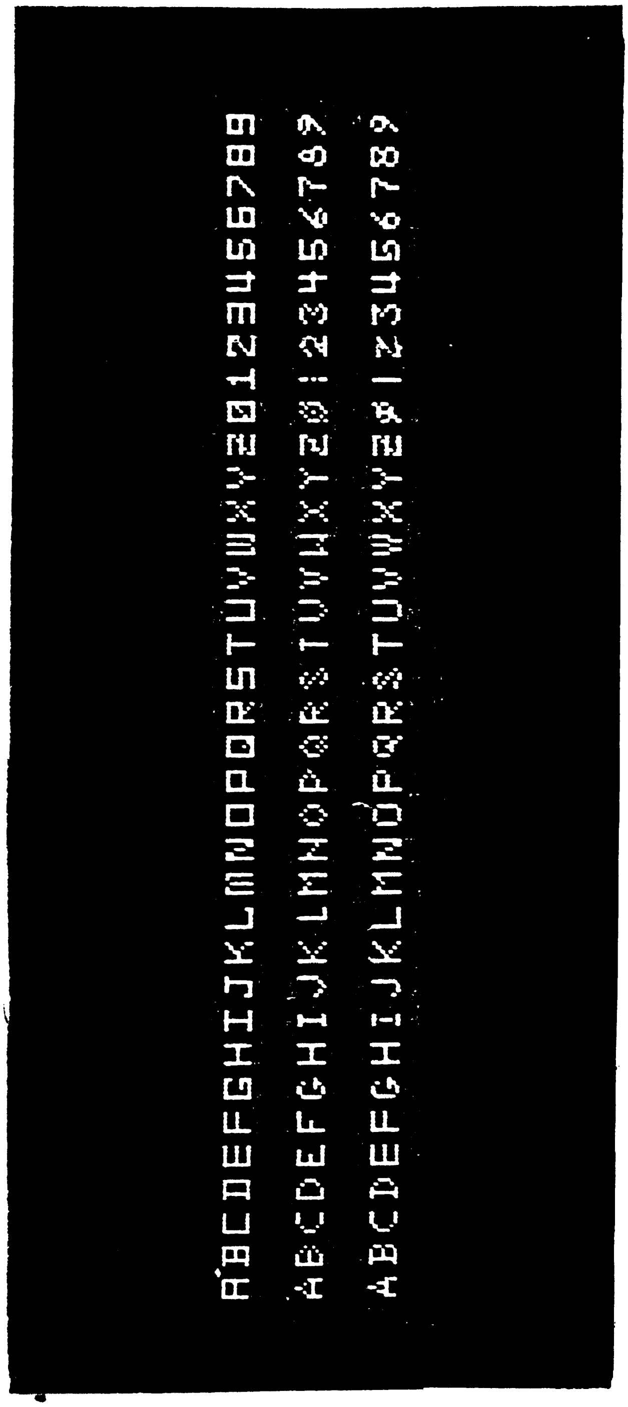


Figure 20. 5x7 Matrix Fonts Examined by Snyder and Maddox (1978).

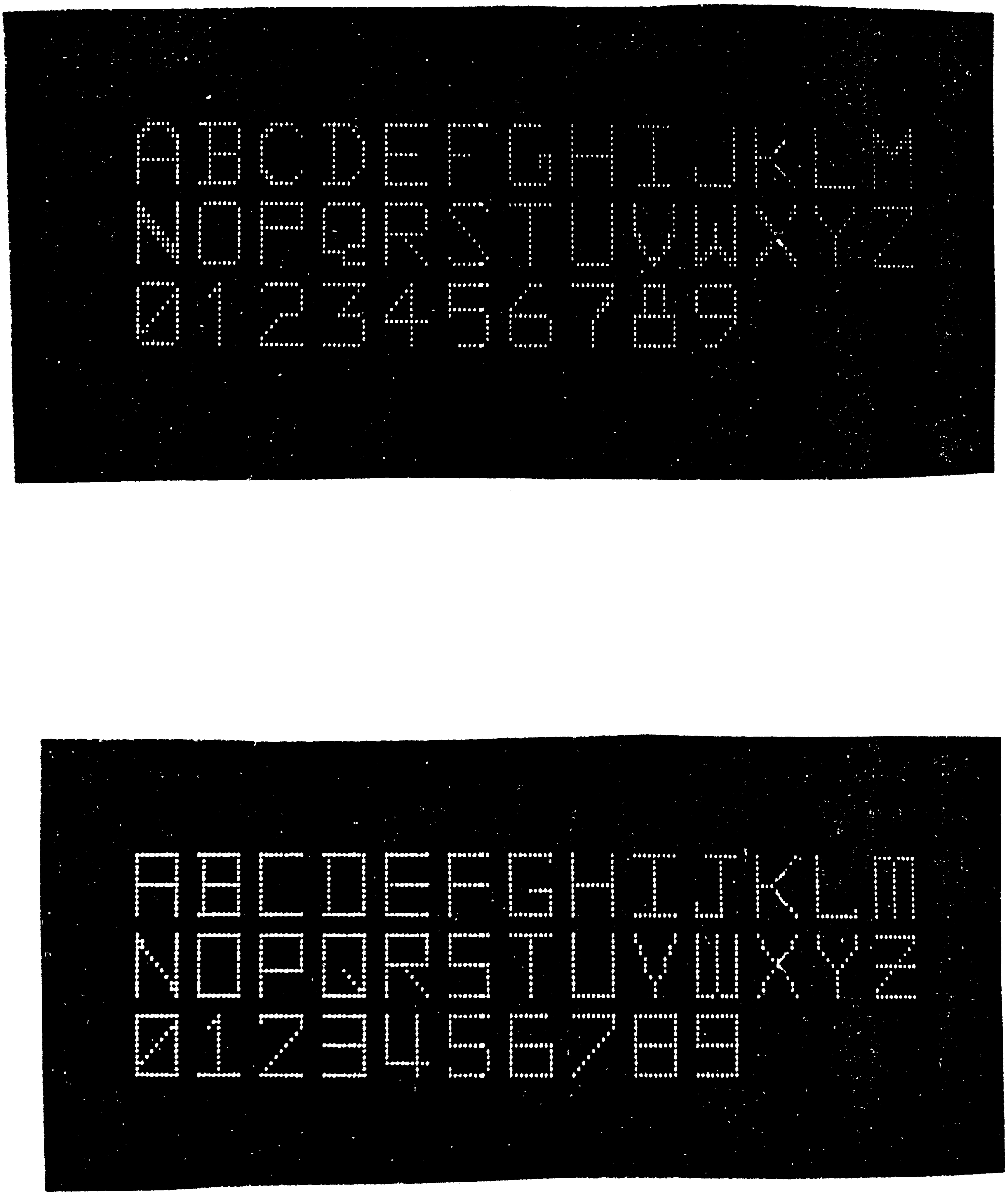


Figure 21. Characters Examined in Multiple Matrix Size Experiment (Snyder and Maddox, 1978).

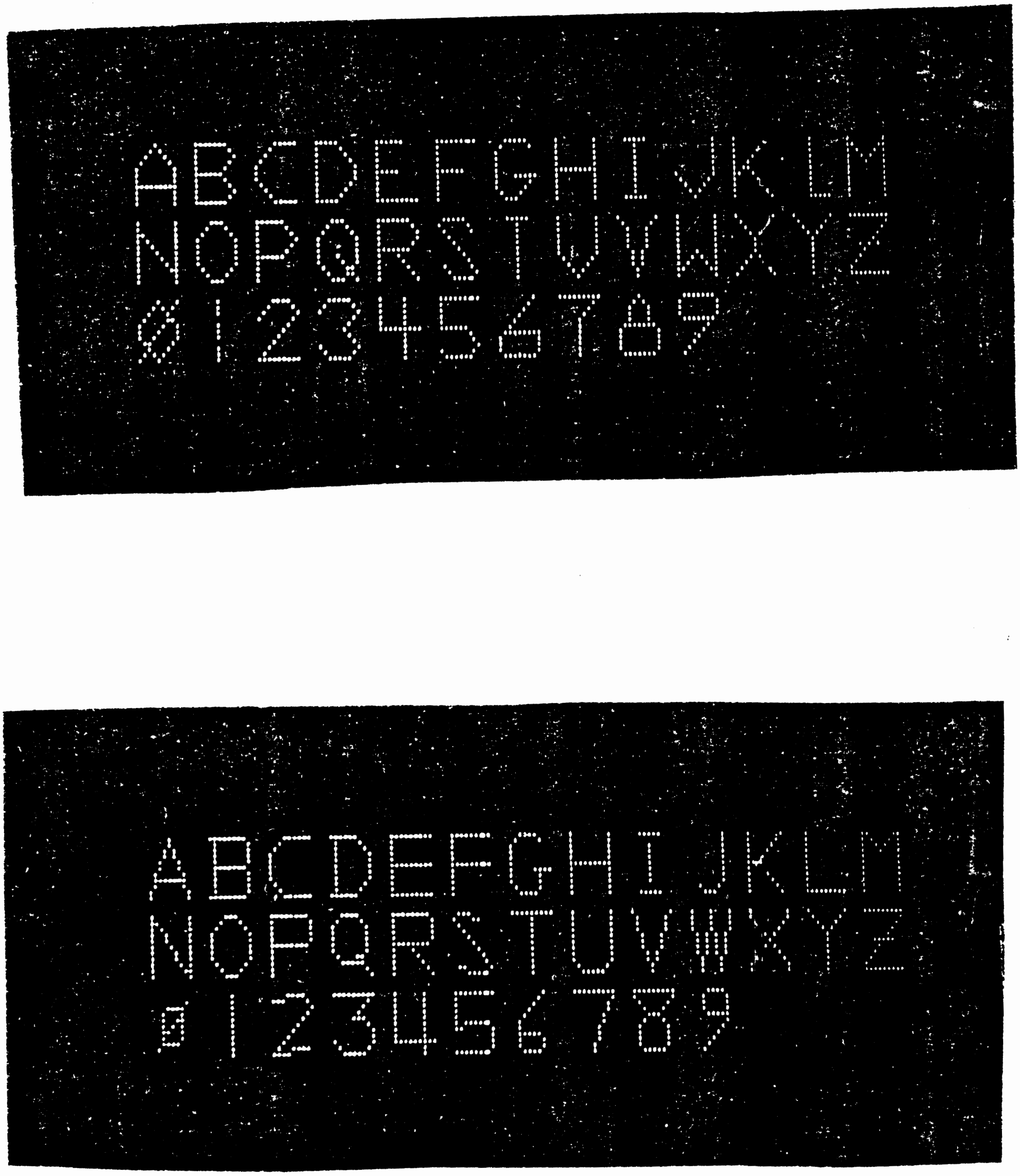


Figure 21. Characters Examined in Multiple Matrix Size Experiment (Snyder and Maddox, 1978).

Characters were viewed from a distance of 102 cm (40 inches) by 40 people. Several character height/dot matrix combinations were explored: 14.4 mm/5x7, 7x9, 9x11; 18.7 mm/7x9; 22.9 mm/9x11. These combinations allowed the character height and matrix size effects to be separated.

Overall, recognition of the Huddleston and Lincoln/Mitre fonts were about equal. Both were superior to the Maximum Dot and Maximum Angle fonts, whose performance was also about equal. With regard to matrix size, increases in the number of dots per character led to reductions in reading errors as did increases in character size. (See Figure 22.)

As can be seen in Figure 22, there were significant interactions between the height/matrix size factor and the font. In particular, there were far fewer errors made in responding to the 5x7 Huddleston font than other fonts for that size. For other heights or matrix sizes, the differences were quite small. Hence, for dot matrix displays, the best font depends on matrix size.

Using the data from all of the above experiments, Snyder et al. carried out additional regression analysis. From that analysis the following models were developed:

Tinker Reading Task

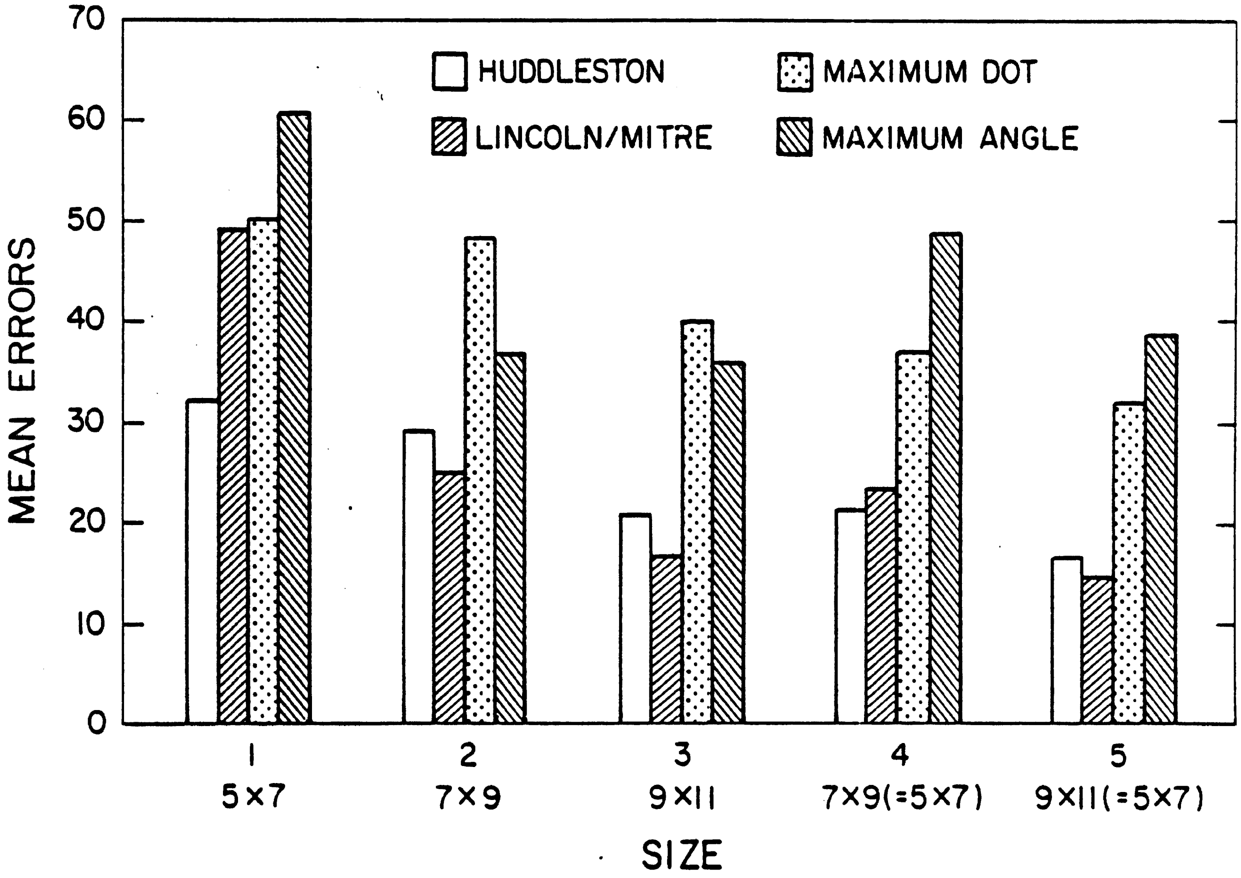
$$\text{Reading Time (s)} = 5.74 + 0.311(\text{HFREQ}) + 2.479(\text{HMOD}) + 4.365(\text{HLOG}) - 14.973(\text{HFLOG}) + 1.112(\text{VMLOG})$$

$$\text{Menu Search Time (s)} = 7.27 + 0.027(\text{HDIV}) + 2.159(\text{HLOG}) + 5.916(\text{VFLOG}) - 0.339(\text{VMTFA}) - 0.054(\text{VRANG}) + 5.487(\text{VMLOG})$$

These equations accounted for about 50% of the variance in search and reading times. The report provides some discussion of the merit and application of these equations, as well as recommendations for the design of dot matrix characters. Readers should bear in mind that these recommendations are primarily for what some would now consider to be low resolution characters. Further, readers should bear in mind that the letters examined were all upper case.

Shurtleff (1980)

This document reviews numerous studies pertaining to the design of 5x7 dot matrix characters. Issues considered include symbol luminance and contrast, horizontal spacing, and dot shape. Of relevance to this section are the studies on font, and a related issue, dot matrix size.



Source: Snyder and Maddox, 1978

Figure 22. Effect of Height/Matrix Size Combinations on the Number of Recognition Errors.

Using an experiment whose procedure was not described, Shurtleff presents data showing how accurately characters in four fonts were recognized. As can be seen from Figure 23, the differences between fonts were small or nonexistent.

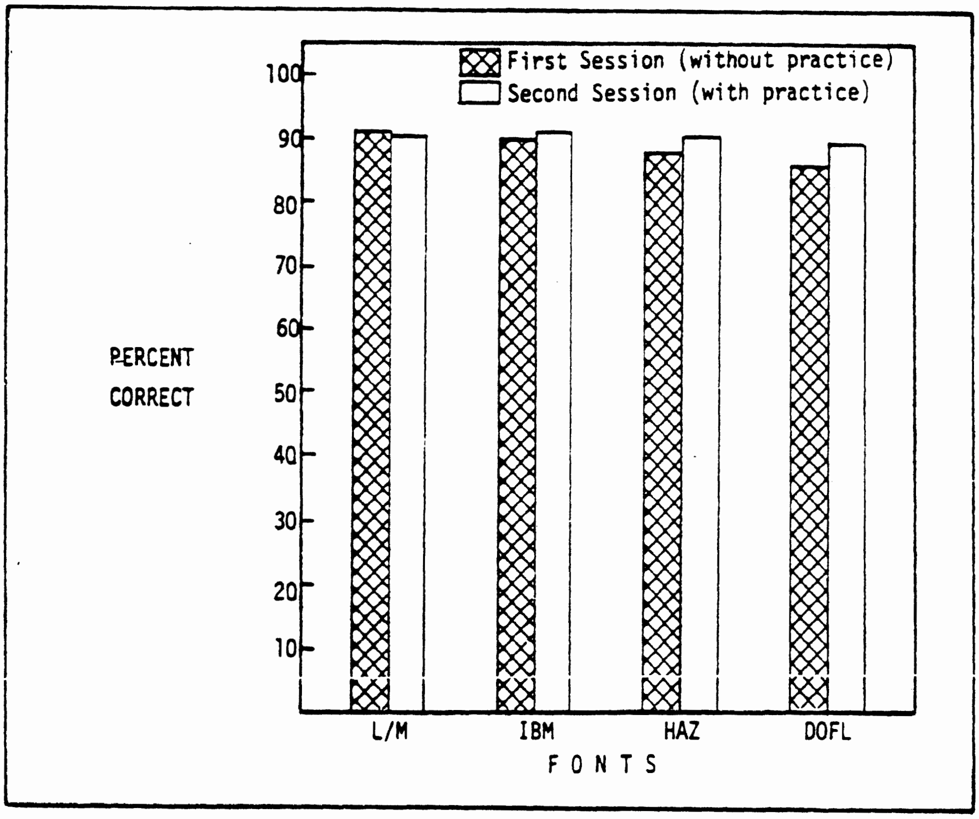
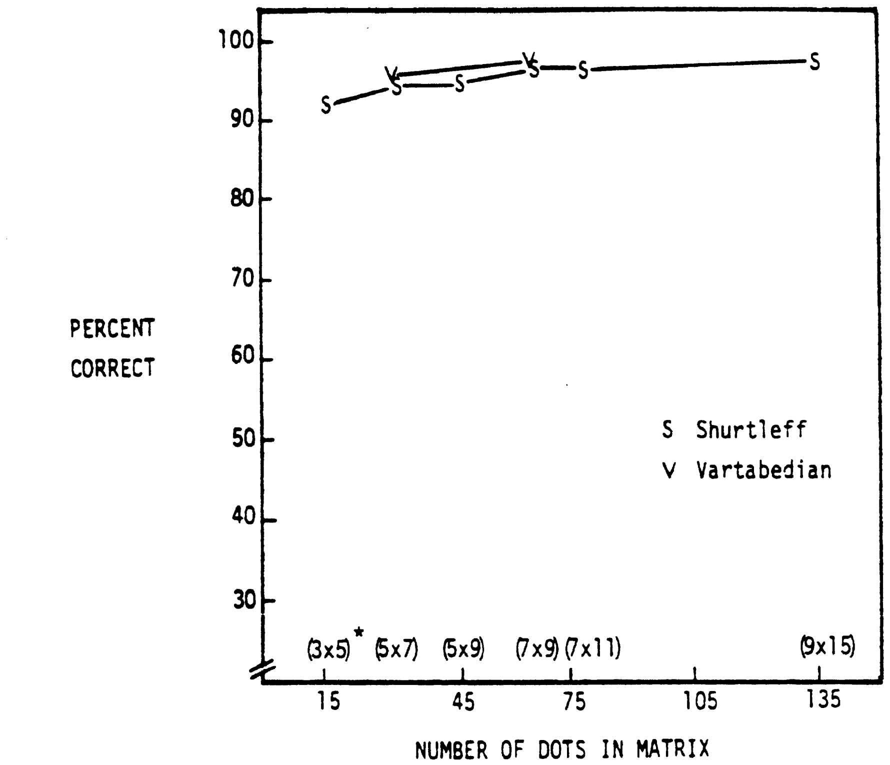


Figure 23. Font Evaluation Described by Shurtleff (1980).

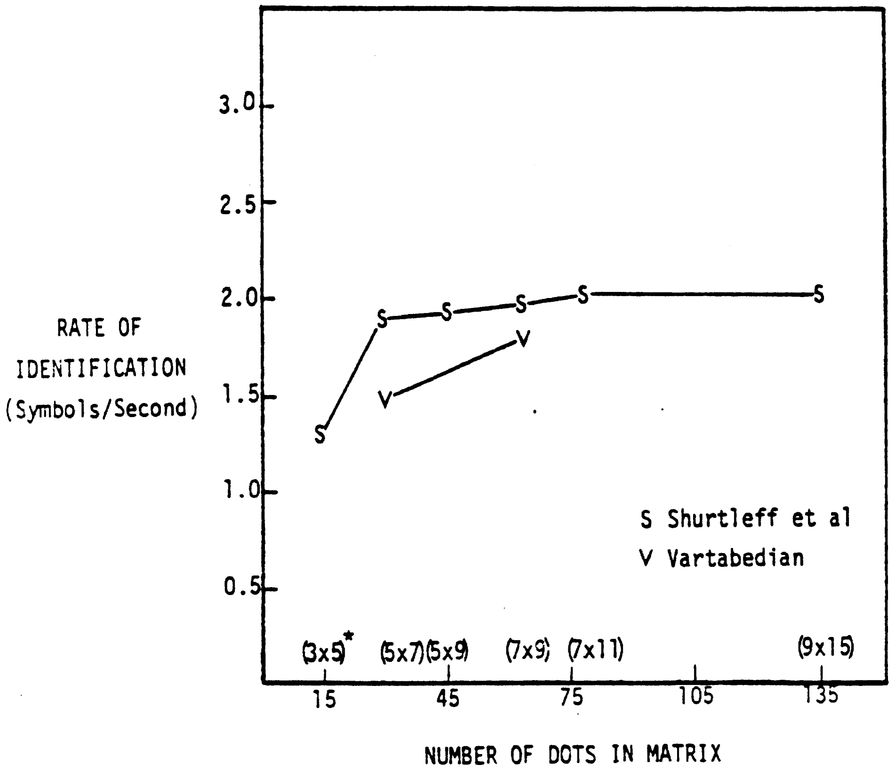
Note: The IBM font refers to the font used on the IBM 029 keypunch. The Hazeltine font is used on its products. The DOFL refers to a font developed by the Diamond Ordnance Fuse Laboratory.

Related to the work on font is work on dot matrix size. Figures 24 and 25 show the results from two studies. Notice that performance gains for matrix sizes larger than 7x11 are negligible.



Source:
Shurtleff (1980)

Figure 24. Relationship Between Dot Matrix Size and Identification Accuracy.



Source:
Shurtleff (1980)

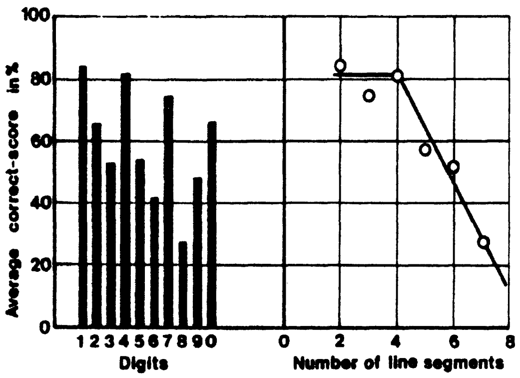
Figure 25. Relationship Between Dot Matrix Size and Rate of Identification.

Van Nes and Bouma (1980)

Van Nes and Bouma (1980) performed experiments on segmented numeral discriminability. Ten people participated in three conditions. Seven-segment displays (19 mm high, 11.5 mm wide, slanted 8 degrees to the right) were used in all conditions. The character luminance was 600 cd/m^2. The illumination and contrast levels were not specified. Since the performance measure was errors, test conditions were chosen to make the percentage of errors large (40%).

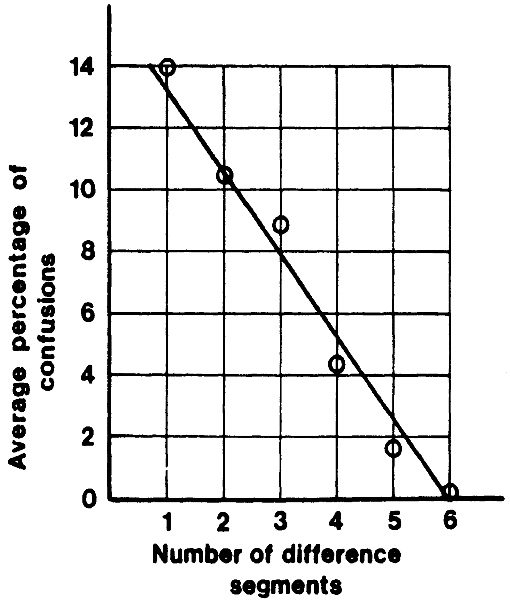
In condition 1, single digits were projected on a screen at a great distance (16 m) and participants said which digit was shown, taking as much time as needed. In condition 2, subjects looked at a fixation point 57 mm ahead. When they pressed a button, a single digit was shown 30 degrees off to the right for 100 ms. In condition 3, there were three digits at eccentricities of 5, 7.5, and 10 degrees. The viewing distance was 57 mm as before. Again, people said which digits appeared.

According to the data, perceptual confusions between the digits shown and the digits identified decreased as the number of segments unique to one number in the pair increased. (See Figures 26 and 27.) Thus, the error rates for the digits 6, 8, 9, and 0 all tend to be high because they are composed of many segments and differ in design by only a single segment. Based on a detailed analysis of the confusion data, the authors developed a procedure for assigning a perceptual weight to each segment, which led to the development of a character set that minimized confusions. Those characters are shown in Figure 28.



Source: Van Nes and Bouma (1980)

Figure 26. Percent Errors for Each Digit Averaged Across Conditions.



Source: Van Nes and Bouma (1980)

Figure 27. Confusion Errors Versus Number of Differing Line Segments.

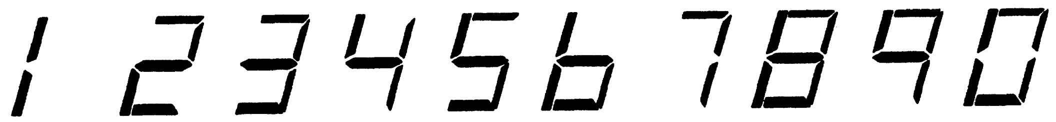


Figure 28. Character Set Proposed by Van Nes and Bouma (1980).

Radl-Koethe and Schubert (1972)

This paper examines the readability of nine different types of displays in production at the time the work was done. The displays did not vary in any consistent manner and hence it is difficult to generalize from the results. Figure 29 shows the designs examined.

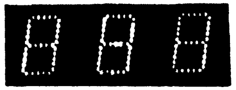
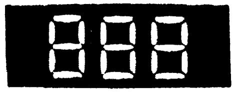
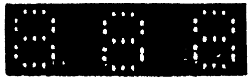
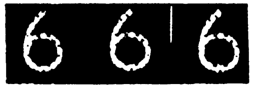
readout system	technology	intensity rise-time constant T_R [msec]	maximal intensity (asb)	letter generation	color	width to height ratio W/H (H = 24')	space between numeral positions
 a)	incandescent bulb	150	12 000	7 segments	white	0,60	1.45 · H
 b)	incandescent bulb	40	16 500	modified 7 segments dot resolved	white	0,55	1.16 · H
 c)	incandescent bulb	150	25 500	7 segments	white	0,60	0.79 · H
 d)	LED (light emitting diodes)	< 1	9	5 x 7 x-y-array	red	0,70	1.46 · H
 e)	LED	< 1	145	7 segments dot resolved	red	0,90	1.69 · H
 f)	Incandescent bulb	200	295	projected arabic numerals	white	0,80	1.38 · H
 g)	glow-discharge tube	< 1	4 200	arabic numerals incandescent filaments	red-orange	0,70	2.09 · H
 h)	glow-discharge tube	< 1	2 300	arabic numerals incandescent filaments	red-orange	0,65	1.15 · H
 i)	incandescent bulb	100	330	arabic numerals dot resolved	white	0,80	1.39 · H

Figure 29. Displays Examined by Radl-Koethe and Schubert.

Displays were mounted in a black panel and illuminated by a 200 lux (18.6 fc) light. The viewing distance was varied between 100 and 215 cm so the visual angle of the display was always 24 minutes of arc. (Displays varied in height from 7 to 15 mm.) Each participant saw several display types nine times at five exposure durations (450 presentations). Their task was to name the digit shown on the display. Because of the experimental design, only 19 of the 40 young men participating saw each type.

Table 3 provides a summary of the results. No statistical tests of significance were provided. However, it is clear from the data that the correlation between reading performance and preferences was not high. In general, people preferred displays showing continuous stroke numerals even though performance in reading them was not very good. (The thresholds for many of the 7-segment and dot matrix displays were less.) Hence, the merits of a particular font or display technique very much depends upon the way it is implemented, depending on such factors as its luminance and contrast ratio.

Table 3. Summary of Results from Radl-Koethe and Schubert.

Display	Threshold (msec)/Rank single #, 50%	Threshold (msec)/Rank 3 numerals, 90% correct	Mis-Readings (%)	Subjective Rank 1=best
a, 7-seg incan	480 (5)	2100 (5)	5.4	6
b, dot incan	140 (2)	290 (2)	5.0	7
c, 7-seg incan	760 (8)	3200 (7)	9.4	8
d, dot LED	70 (1)	240 (1)	8.3	4
e, 7-seg LED	840 (9)	---- (9)	14.8	9
f, stroke incan	730 (7)	1070 (4)	1.5	3
g, stroke incan	380 (4)	>1000 (8)	11.5	2
h, stroke incan	170 (3)	>2700 (6)	5.5	1
i, stroke incan	600 (6)	940 (3)	3.7	5

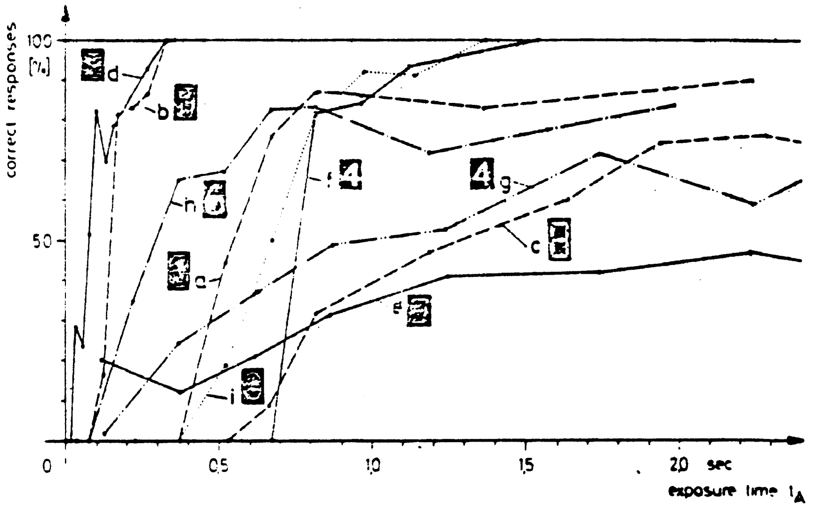
Figure 30 is a detailed presentation of the legibility thresholds. Notice from this figure that the relative ordering of display types varies considerably with the exposure durations and task. (The curves cross.)

Summary

A key finding from the research on font (e.g., Hind, Tritt, and Hoffman, 1976) is that its effect on performance (relative to changes in contrast and illumination) is secondary. Further, relative minor changes in font (for example, slant as examined by Plauth, 1970) have virtually no effect on performance.

With regard to the kinds of fonts that are "best," the Snyder and Maddox (1978) work shows that both the Huddleston and

Lincoln/Mitre dot matrix fonts are superior to other alternatives although the merits of a particular font depend upon the dot matrix font used. Clearly, dots should be closely spaced and round. Both that study and the Shurtleff work show there are merits (in terms of performance) in using matrices of greater resolution than 7x9 (e.g., 7x11). Performance with high resolution matrices (e.g., beyond 9x13) has generally not been explored. In selecting character sets, the emphasis has been on performance. There may be aesthetic reasons for selecting higher resolution.



Source: Radl-Koethe and Schubert (1972)

Figure 30. Legibility of Three Numerals As a Function of Exposure Time.

All this, however, does not mean one is free to choose any font. For example, Smith (1978) examined problems people had in reading the Lansdell font, a highly distinctive set designed to minimize discriminability problems. He found that because the design was unusual, people had problems reading that font.

The discriminability issue is also examined by Van Nes and Bouma. Their work led to a predictive model for the discriminability of seven-segment characters, the type currently used for numeric automobile speedometers. In that model, confusion errors were proportional to the number of segments in which the digit pair of interest differed. This suggests that making some of the segments slightly different (with the model suggesting which ones) could enhance the discriminability of these characters. For example, the discriminability of the digits "5" and "6" could be increased by modifying the shape of the lower left vertical segment.

EFFECT OF COLOR ON LEGIBILITY

Introduction

While the mechanisms of human color vision are well understood and there are numerous studies in which the legibility of color displays has been examined, there are only a few studies that have attempted to quantify chromatic (color) contrast.

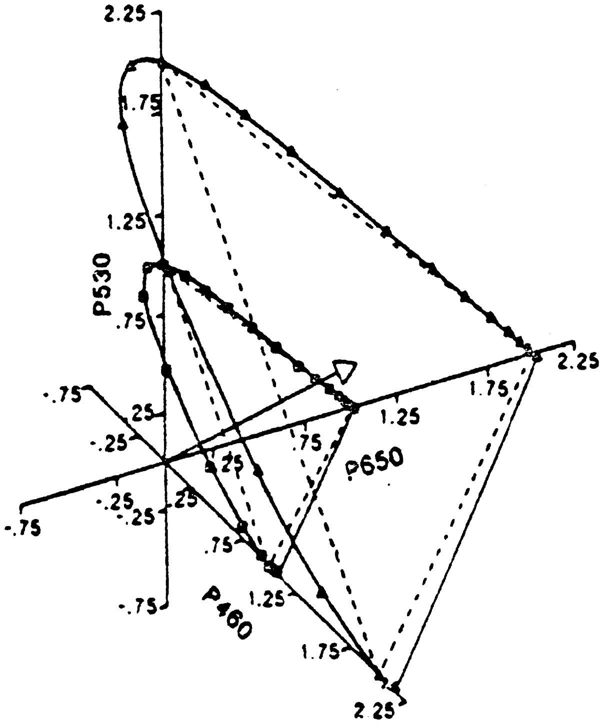
Critical to understanding the research on this subject is some knowledge of the measurement of color and color contrast. Those familiar with this subject should skip to the next subheading.

The colors that people perceive are a direct result of human physiology. People have two types of light sensing cells, rods and cones. There are three types of cones, each of which is sensitive to a different portion of the visible spectrum (400-700 nm). In reference to the portions of the spectrum to which they are sensitive, the cones are referred to as red, green, and blue. Hence, the particular color a person sees depends upon the relative activity of these three cone systems. Further, if one carefully chooses the spectral distributions of lights, there are multiple combinations which will all seem to have the same color.

In the most commonly used system for specifying color, the CIE system, this fact is the cornerstone of the specification procedure. One can think of the color space as a three-dimensional coordinate system with each axis representing the contribution of three standard lights (red, green, and blue). If the axes are scaled so that the distance on each axis represents the relative contribution (intensity) of each of the three arbitrary lights, then Figure 31 results.

In that figure, the plane where $x + y + z = 1$ is used to map out the colors people can see. For convenience, this figure is often presented showing the projection of this surface in the x-y plane (Figure 32, the 1931 CIE standard observer).

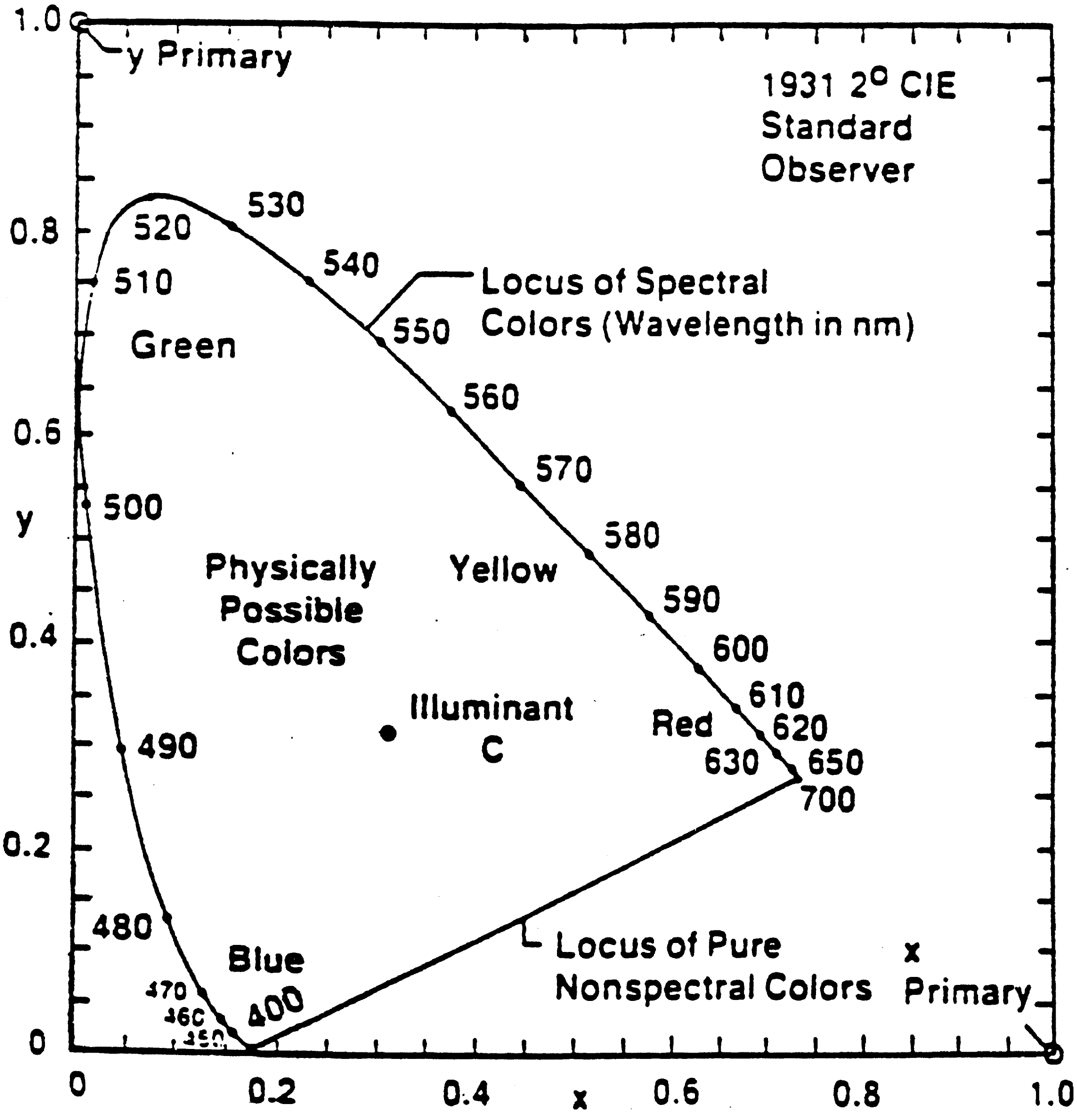
Several studies have been carried out to determine the minimum color difference that people can distinguish (i.e., the color contrast threshold). In those studies a small number of observers (often one) are shown circular color patches which are split vertically into two halves that differ slightly in color. Observers use a control to adjust the image so the two halves are just noticeably different. Shown in Figure 33 are some typical results. The ellipses represent how far one must go from specific points in the CIE space to observe a just noticeable difference.



Source: Boff and Lincoln (1988)

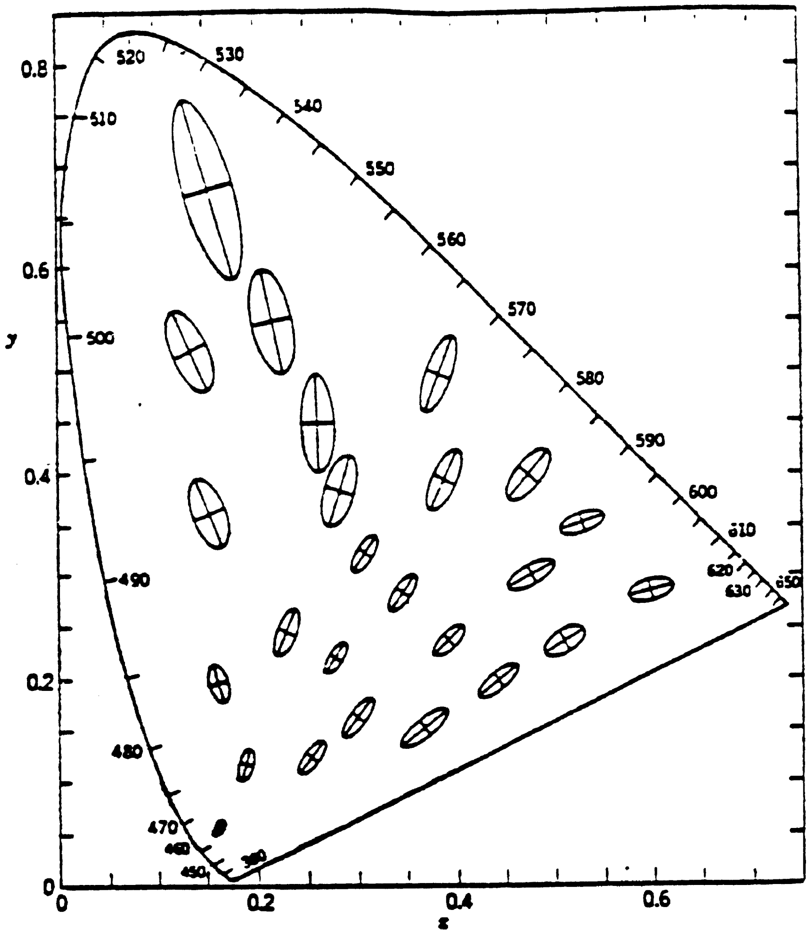
Figure 31. Three Axis Coordinate System.

It is generally assumed that the difference between two colors, their chromatic contrast, is proportional to how many just noticeable differences (jnds) separate their coordinates on the CIE diagram. The problem with using Figure 33 is that the relationship between physical distance in the CIE diagram and perceptual distance is not linear. Hence, a number of efforts have been made to rescale the CIE diagram to make those two quantities equal. In a physical sense, this means making all of the ellipses in Figure 33 (the measure of perceptual differences) equal in all directions (round) and in size (diameter). Shown in Figure 34 is the result of such a rescaling effort.



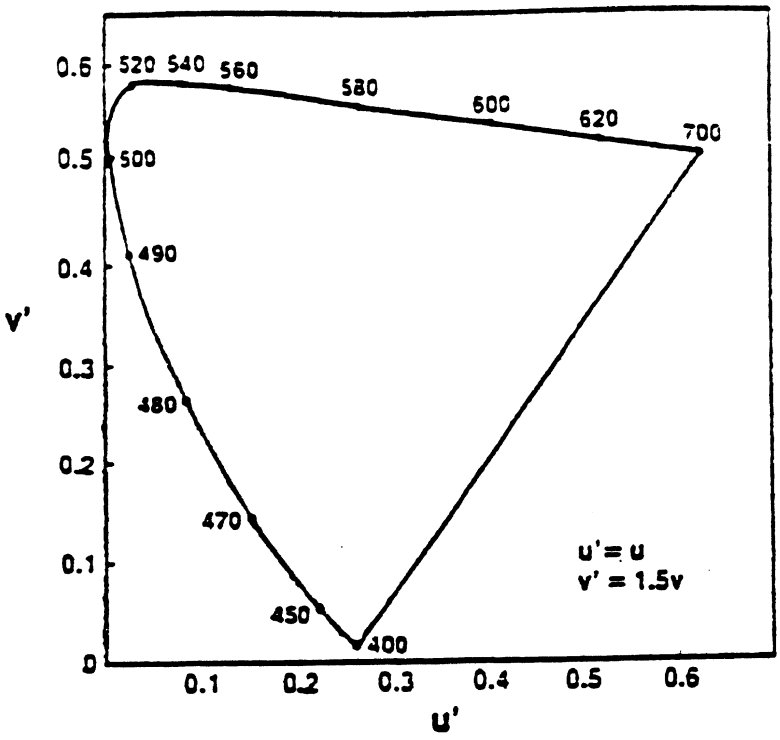
Source: Silverstein and Merrifield (1985)

Figure 32. The 1931 CIE Standard Observer.



Source:
Silverstein and
Merrifield (1985)

Figure 33. MacAdam Ellipses.



Source:
Silverstein and
Merrifield (1985)

$$u' = \frac{4x}{-2x+12y+3}$$

$$v' = \frac{9y}{-2x+12y+3}$$

Figure 34. CIE 1976 u'v' Chromaticity Diagram.

Again, in this diagram, equal distances between points represent equal differences in appearance. To compute how different two items appear (e.g., the contrast between a character and a background), it is usually assumed that the total difference in appearance is the sum of the number of jnds in luminance contrast and the number of jnds in chromatic contrast. According to Carter and Carter (1983), that measure (Delta_E) is computed as follows using the transforms between the CIE L*u*v* (CIELUV) coordinates and the location in u'v' space:

$$\text{Delta\_E} = [\ (\text{Delta\_L*})^2 + (\text{Delta\_u*})^2 + (\text{Delta\_v*})^2\ ]^{1/2}$$

where:

$$L^* = 116(Y/Y0)^{1/3} - 16,\ Y/Y0 > .01$$
$$u^* = 13L^*(u' - u0')$$
$$v^* = 13L^*(v' - v0')$$
$$u' = 4X/(X + 15Y + 3Z)$$
$$v' = 9Y/(X + 15Y + 3Z)$$

Further, Carter and Carter note that for surface colors, Y0 is the Y tristimulus value of the reference white for the selected illuminant and observer. For self-luminous displays, Y0 is the maximum possible luminance of the images whose difference is to be calculated. Those interested in further details of the calculations should retrieve the Carter and Carter (1983) paper. Discussions of chromatic contrast also appear in Kuehni (1982) and Boynton, Nagy, and Olson (1983). Particularly clear and comprehensive discussions of the general topic of chromatic contrast appear in Billmeyer and Saltzman (1981) and Silverstein and Merrifield (1985).

In the remainder of this section, several studies that try to assess the equivalence of luminance and chromatic contrast are described. Considerable research on this topic has been conducted in the last few years and the studies presented are only a sampling of the more relevant current work.

Post, Costanza, and Lippert (1982)

This paper reports two experiments comparing the relationship between achromatic contrast and chromatic contrast. In the initial experiment, six people adjusted the luminances of pairs of seven colored patches on a video display until their luminance difference matched the difference of an adjacent achromatic pair. Three luminance levels were examined. The 63 combinations (7 x 3 x 3) were shown to each person once in each of 4 sessions in a random order.

Several regression analyses were carried out to predict luminance modulation, a measure of achromatic contrast, from the data. In particular:

$$Lmod^2 = .0226 + 3.9778(dL^2) + .135(du^2) + .065(dv^2)$$

where:

$$Lmod = (Lmax - Lmin)/(Lmax + Lmin)$$

A second experiment was conducted to examine stimuli that varied in brightness, as well as hue and saturation. Each of 3 people saw 210 combinations of those 3 lighting variables. The same general test procedure was used. In this case, the regression analysis indicated:

$$Lmod^2 = .1053 + 2.7746(dL^2) + .296(du^2) + .0914(dv^2)$$

Kimura, Sugiura, Hiroaki, and Nagai (1988)

This article discusses the legibility of CRT's as used in automobiles. A series of experiments was conducted to examine the effects of chromaticity difference, luminance contrast and background luminance on legibility and comfort.

The first experiment examined the effect of character-background color combination on visual recognition time. Subjects were seated in a dark room 31.5 in (80 cm) from a 10-inch (25.4 cm) color CRT. A total of 240 character-background color combinations were examined. A four-character Chinese word was presented on the CRT. If the observer was unable to read the word, they pressed the space bar so the word would remain on the screen for another brief, but unspecified, period of time. This process was repeated until the word was readable. The word and its position on the screen changed with each color combination.

The experimenters found that as luminance contrast and chromaticity difference approached zero, recognition time increased. A family of equal recognition time curves, shown in Figure 35, was developed. Functions of this type can be used to evaluate tradeoffs between luminance and chromatic contrast. The contours would have been much more useful if the times were labeled in Figure 35.

Source:
Kimura, et al.
(1988)

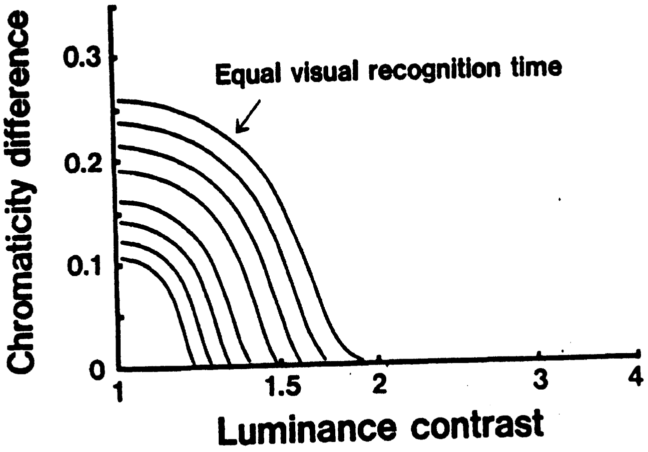


Figure 35. Recognition Time Contours.

Note: The chromaticity difference is the distance between the character and background coordinates in the CIE 1976 UCS Chromaticity Diagram. The luminance contrast is the ratio of the character luminance to the background luminance.

The second experiment addressed the issue of visual comfort. Drivers viewed two unspecified targets, one at the center of the console, and one at the center of the instrument cluster of a vehicle. While driving a vehicle on rural expressways or urban streets, participants kept their eyes on the targets until they were "uncomfortable." It is not clear if discomfort was due to the glare of the targets or the lack of information about the road scene. Fixation times were determined from electro-oculograms (EOGs). Results are presented in terms of the probability that a driver can view a target for a given period of time without becoming uncomfortable. For example, the probability of viewing a target for 2.0 sec without discomfort is 50% (1.0 sec, 95%; 0.8 sec, 99%).

Kimura et al. also include a figure showing "the result of the subjective evaluation of previous CRTs for automobiles" (Kimura, Surhuguia, Shinkai, and Nagai, 1988, p. 3). Further details are not provided. This figure (Figure 36 here) suggests that for achromatic displays to be "easy to read," luminance contrast should be greater than 1.5. For chromatic displays where luminance contrast is absent, the chromaticity difference should be greater than 0.2. It is unknown for what illumination conditions these suggestions apply or how "easy to read" was defined.

Source:
Kimura et al. (1988)

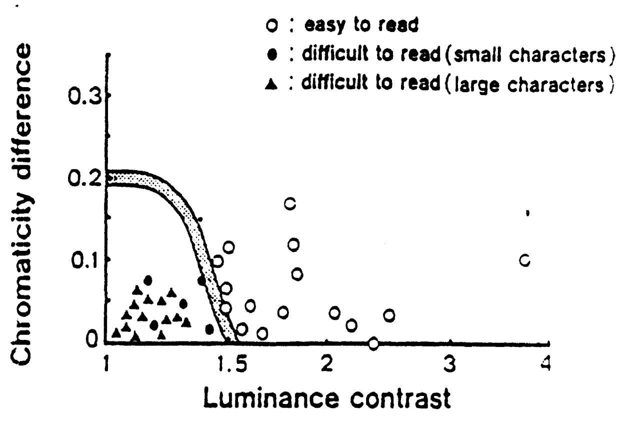


Figure 36. Subjective Evaluations - Legibility of Automotive CRTs.

Following the discomfort analysis, the authors provide a series of very involved calculations to determine appropriate levels of chromaticity difference and luminance contrast. This difference is the square root of the sum of the squared differences of the u', v' coordinates. (This expression is simply the Pythagorean Theorem. The distance between two points in space (the hypotenuse of a right triangle) is the square root of the sum of their squares.) As seen in the figure, calculated values are relatively close to measured values, although statistical tests comparing the two are not provided. It is not clear from the paper how the authors measured chromaticity.

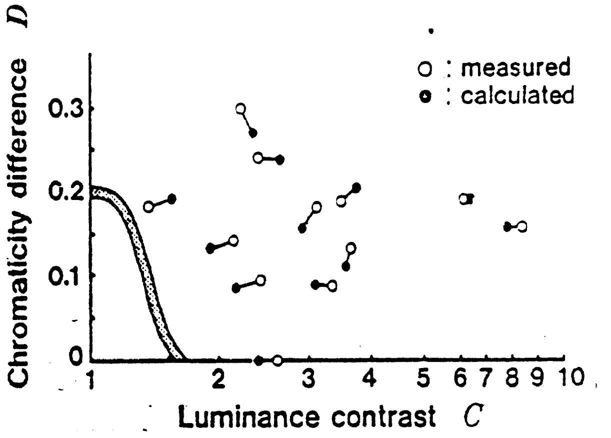


Figure 37. Measured vs. Calculated Contrasts in Kimura et al. (1988).

Silverstein, Lepkowski, Carter, and Carter (1986)

This paper describes how to select colors for visual displays and provides an algorithm for determining "near-optimal" color sets. The process of color selection becomes especially difficult when any of the following conditions exist: (1) color displays are to be used in dynamic, high-ambient lighting environments; (2) a large number of colors are required for information coding; (3) color display media with a restricted color gamut are to be used; (4) high information density and small color image sizes are required; (5) observers with defective color vision are potential display users; and (6) operator performance is critically linked to accurate color discrimination or color identification.

Table 4 depicts qualitatively the principal factors affecting the ability to distinguish between display colors.

Based on their work, Silverstein et al. developed a computer program to determine near optimal color sets. It incorporates the following features: (1) internal colorimetric modifications to display colors for user-specified display background luminance and chromaticity (to adjust for reflected ambient light); (2) automatic establishment of display background luminance and chromaticity when one or more colors in the selection process have fixed values; (3) a correction factor which estimates the biases in color perception for small color image field sizes; and (4) a correction factor which estimates the biases in color perception for observers with red/green color vision deficiencies. Program output includes the names of all fixed and variable colors, their chromaticity coordinates, their RGB luminance values, Delta_E for all pairs of colors, and several other measures. Readers are encouraged to examine the technical reports on which this work is based for further details (Silverstein and Merrifield, 1985; Silverstein, 1987).

Table 4. Principal Factors Affecting the Ability to Distinguish Between Display Colors.

	Change in Factor	Ability to Distinguish Colors
Wavelength separation	increase	increase
Color purity	increase	increase
Brightness	increase	increase
Color stimulus size	increase	increase
Brightness adaptation level	increase	increase
Number of colors	increase	decrease
Display background:		
Light	N/A	increase
Dark	N/A	decrease
Color stimulus location:		
Central	N/A	increase
Peripheral	N/A	decrease
Type of discrimination required:		
Relative/comparative	N/A	increase
Absolute-identification	N/A	decrease
User population characteristics:		
Age	increase	decrease
Color vision anomalies	N/A	decrease

Human Factors Society (1988)

This document is the U.S. national standard (ANSI, American National Standards Institute) for the design of computer workstations. While it is intended for office applications, it is likely to be applied to other contexts. Sections 6.17 and 9.7 of this standard address the issues of symbol color contrast and legibility. To provide for adequate legibility of colored symbols on a colored background, the standard states that Delta_E (CIE Yu'v' distance), not Delta_E (CIE Lu*v*), should exceed 100. For text, lower values sometimes provide adequate levels of readability.

Using the 1976 CIE UCS color diagram, Delta_E is computed as follows:

$$\text{Delta\_E}(Yu'v') = [(155\ TB/M)^2 + (367\ U)^2 + (167\ V)^2]^{0.5}$$

where TB = difference in luminance between text and background
M = maximum luminance of text or background
U = difference between u' coordinates of text and background
V = difference between v' coordinates of text and background

The standard also includes minimum requirements for character height and luminance contrast.

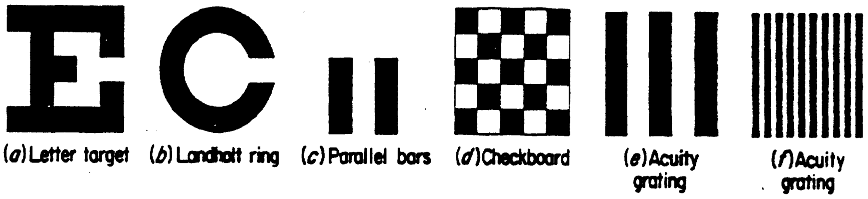
Summary

Each of the documents reviewed here assumes that luminance and chromatic contrast can be added together to compute total contrast, Delta_E. In both cases, the measure of contrast is the number of jnds between the levels of interest. The use of the Delta_E measure in design is just beginning to be accepted practice, though how and when it should be used is still open to discussion. Readers interested in color are strongly encouraged to read the Silverstein and Merrifield technical report. Further, automobile designers are strongly encouraged to look at the luminance and Delta_E requirements in the ANSI standard. While the minimums in that standard are reasonable for office applications, they are probably too low for automotive applications. Unlike reading text in an office, driving and reading vehicle displays is a more complex, time-sharing task. Furthermore, increases in reading time for vehicle displays can have critical consequences to drivers, outcomes for which there are few parallels in an office.

ADULT VISUAL CAPABILITIES

What Is Visual Acuity?

In order to design a display for drivers to see, one must know their visual acuity. Visual acuity is "the ability to resolve (distinguish) black and white detail" (McCormick, 1970, p. 59). Generally, it refers to being able to see a small detail with a high degree of accuracy. Figure 38 shows some of the materials used for acuity tests. In applied work, the "E" and Landolt C are most commonly used.



Source: McCormick (1970)

Figure 38. Materials Used in Tests of Visual Acuity.

Acuity is usually expressed as a ratio (the Snellen Ratio) of two numbers such as 20/40. The term in the denominator refers to how far away from a target (here 40 feet) the standard observer must be to see what a person with normal vision can see at the distance in the numerator (20 feet). The standard target is a 1 minute of arc critical detail. In this example, critical details have to be twice the size required by a person with good vision. Should a 6 appear in the numerator (e.g., 6/12), the distances are in meters. Should a 14 appear in the numerator, the expression refers to near acuity (at 14 inches).

Roberts (1964)

There are two key recent studies on acuity: the 1960 U.S. government study carried out as part of the Health Examination Survey, and a survey of English drivers. In the U.S. survey (Roberts, 1964) a national probability sample of 6,672 people were examined. Both near acuity (14 inches) and far acuity (20 feet), "corrected" and "uncorrected" (without glasses or contacts) were examined using a commercial instrument, the Sight Screener. Participants reported the letters shown. Both right eye, left eye, and binocular acuity were recorded. For instrument panel design, the near acuity data are more pertinent.

Shown in Table 5 are the uncorrected acuity data for men and women. Table 6 presents similar data for corrected acuity.

Averaged across age groups, about 24.0% of the population has better than normal acuity (14/14) uncorrected (29.7% when corrected), 20.7% have normal acuity (35.2% corrected) and the remainder is worse than normal.

Tables 7 and 8 contain the more commonly reported statistic, far acuity. In comparing the near and far acuity distributions by matching acuity ratios (e.g., 14/14 with 20/20), far acuities tend to be lower. For example, 38.4% of the adults have 20/15 acuity whereas only 28.6% have 14/10.5. Similarly, only 2.8% have a far acuity 20/100 or worse, but 4.4% have a near acuity of 14/70 or worse.

In the case of vehicles, people are screened before they are licensed, so those with extremely poor vision (e.g., so the 1.4% with 14/140 or worse corrected) are not allowed to drive. In most states the additional 3% with 14/70 vision (the near acuity parallel of 20/100) are also not allowed to drive. However, screenings only happen every several years, and while some people are required to wear glasses when they drive, many don't.

When designing a display, the selection of the design acuity level is an important decision. Typically, human factors people use percentiles when designing systems. The 95th percentile (95% of the population within the design limits) is commonly chosen for single anthropometric dimensions. For the remaining 5%, it is assumed some sort of adjustment will be provided. However, since one certainly wouldn't design a display that only worked 95% of the time, it doesn't make sense to design a display that only 95% of drivers can just barely read. Displays should be easy to read. When designing something (e.g., a bridge, an aircraft wing), one typically computes the design limit and then multiples that number by a safety factor (anywhere from 2-18), and uses the resulting value in the actual design.

Davison and Irving (1980)

Davison and Irving (1980) report the distribution of binocular visual acuity for 1,368 drivers of all ages. The sample age distribution was matched to the driving population. The data was collected at 25 sites in the U.K. using a Snellen-E test chart. People who drove to the test site wearing glasses or contact lenses wore them during the test.

Table 5. Percentage of Adults Reaching Specified Acuity Levels-Uncorrected Near Vision.

Sex and acuity level	Total, 18-79 years	18-24 years	25-34 years	35-44 years	45-54 years	55-64 years	65-74 years	75-79 years
Both sexes	Percent distribution							
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
14/7 or better	1.0	1.7	2.1	1.4	0.1	-	-	-
14/10.5	23.0	48.4	46.9	31.4	1.9	0.1	-	-
14/14	20.7	33.2	34.8	35.9	7.5	0.7	-	3.5
14/21	8.9	9.6	8.3	14.5	11.3	3.1	2.2	1.2
14/28	4.7	2.6	1.5	4.4	9.2	5.7	4.3	7.7
14/35	4.4	1.0	1.0	2.2	8.7	7.1	7.7	7.4
14/49	5.5	0.4	1.1	2.2	12.5	10.8	7.2	8.5
14/70	15.7	2.1	1.7	5.4	29.6	32.6	30.8	33.8
14/140	11.8	0.6	1.9	2.0	15.3	31.2	31.2	22.3
Less than 14/140	4.3	0.4	0.7	0.6	3.9	8.7	16.6	15.6
Men								
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
14/7 or better	1.6	2.6	3.6	2.3	0.2	-	-	-
14/10.5	26.9	56.2	53.7	38.4	2.6	0.2	-	-
14/14	19.1	28.5	29.2	34.4	9.4	0.8	-	7.0
14/21	7.7	7.4	4.9	12.6	12.2	3.0	2.4	1.1
14/28	4.5	1.7	1.9	3.5	8.3	5.4	4.3	14.0
14/35	4.4	0.4	0.9	2.2	8.2	8.4	8.1	6.6
14/49	6.2	0.4	1.5	1.7	12.7	12.8	10.0	10.3
14/70	16.3	1.9	1.8	3.6	31.0	35.2	33.4	32.6
14/140	9.9	0.5	2.1	0.8	13.0	26.9	27.7	14.6
Less than 14/140	3.4	0.4	0.4	0.5	2.4	7.3	14.1	13.8
Women								
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
14/7 or better	0.4	0.8	0.8	0.6	-	-	-	-
14/10.5	19.6	41.7	40.7	25.1	1.3	-	-	-
14/14	22.1	37.1	39.7	37.1	5.7	0.7	-	-
14/21	9.9	11.5	11.4	16.2	10.5	3.2	2.0	1.3
14/28	4.9	3.4	1.2	5.1	10.0	5.9	4.3	1.5
14/35	4.4	1.6	1.1	2.3	9.2	5.8	7.4	8.3
14/49	5.0	0.5	0.8	2.7	12.4	9.1	4.9	6.6
14/70	15.2	2.3	1.5	7.0	28.0	30.2	28.7	35.2
14/140	13.4	0.7	1.8	3.2	17.5	35.0	34.1	29.8
Less than 14/140	5.1	0.4	1.0	0.7	5.4	10.1	18.6	17.3

Table 6. Percentage of Adults Reaching Specified Acuity Levels-Corrected Near Vision.

Sex and acuity level	Total, 18-79 years	18-24 years	25-34 years	35-44 years	45-54 years	55-64 years	65-74 years	75-79 years
Both sexes	Percent distribution							
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
14/7 or better	1.1	1.3	2.4	1.7	0.5	0.2	-	-
14/10.5	28.6	53.4	53.1	37.3	9.2	5.5	2.3	-
14/14	35.2	36.2	36.4	43.0	34.1	33.5	23.8	13.0
14/21	19.8	7.7	6.4	12.2	29.7	33.2	40.3	30.0
14/28	6.2	0.8	0.6	2.8	10.4	11.2	13.4	20.7
14/35	2.8	0.3	0.3	1.1	5.0	5.4	5.8	9.4
14/49	1.9	0.1	0.2	0.6	3.8	2.8	3.5	9.8
14/70	3.0	0.1	0.4	1.0	5.8	5.0	6.7	11.6
14/140	1.0	0.1	0.2	0.1	1.1	2.6	3.2	1.7
Less than 14/140	0.4	-	-	0.2	0.4	0.6	1.0	3.8
Men								
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
14/7 or better	1.9	2.3	4.2	2.6	0.8	0.2	-	-
14/10.5	33.0	59.9	60.0	43.9	11.2	7.4	3.2	-
14/14	32.2	29.3	29.6	39.0	35.0	32.9	24.0	17.0
14/21	17.2	7.4	3.9	11.0	26.0	27.5	37.6	26.4
14/28	5.5	0.7	1.2	1.4	8.2	10.4	13.4	22.4
14/35	3.2	0.4	0.4	0.6	5.0	7.6	7.2	11.2
14/49	2.2	-	0.2	0.4	5.3	3.6	3.4	9.1
14/70	3.4	-	0.4	0.7	7.1	6.4	7.2	9.3
14/140	1.1	-	0.1	0.1	1.3	3.6	3.2	1.5
Less than 14/140	0.3	-	-	0.3	0.1	0.4	0.8	3.1
Women								
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
14/7 or better	0.5	0.6	0.8	0.8	0.3	0.2	-	-
14/10.5	24.7	48.0	46.7	31.3	7.2	3.7	1.5	-
14/14	37.8	41.9	42.5	46.8	33.3	34.2	23.5	9.0
14/21	22.2	7.9	8.7	13.3	33.3	38.4	42.4	33.5
14/28	6.7	0.8	0.1	4.1	12.4	11.8	13.4	18.9
14/35	2.5	0.3	0.3	1.5	5.0	3.3	4.7	7.8
14/49	1.5	0.2	0.1	0.7	2.4	2.1	3.7	10.4
14/70	2.7	0.1	0.5	1.3	4.5	3.7	6.4	13.9
14/140	0.9	0.2	0.3	0.1	0.9	1.8	3.2	1.9
Less than 14/140	0.5	-	-	0.1	0.7	0.8	1.2	4.6

Table 7. Percentage of Adults Reaching Specified Acuity Levels-Uncorrected Far Acuity.

Sex and acuity level	Total, 18-79 years	18-24 years	25-34 years	35-44 years	45-54 years	55-64 years	65-74 years	75-79 years
Both sexes	Percent distribution							
Total-------------	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
20/10 or better---------	1.1	1.8	2.1	1.5	0.7	-	-	-
20/15------------------	29.2	48.3	48.0	42.3	17.5	4.2	0.9	-
20/20------------------	23.6	24.7	27.2	32.9	26.4	16.8	4.8	1.5
20/30------------------	15.4	10.4	9.6	10.9	23.7	21.8	19.3	13.1
20/40------------------	6.5	3.6	2.5	3.1	7.8	11.8	13.3	17.6
20/50------------------	4.6	1.9	1.5	1.9	6.3	8.3	9.9	12.3
20/70------------------	3.5	1.0	1.6	1.5	3.8	7.1	6.9	13.5
20/100-----------------	9.6	3.6	2.8	2.8	8.5	20.2	28.9	28.0
20/200-----------------	4.1	3.0	2.4	1.9	3.8	7.7	9.2	6.0
Less than 20/200--------	2.4	1.7	2.3	1.2	1.5	2.1	6.8	8.0
Men								
Total-------------	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
20/10 or better---------	1.4	1.8	3.2	1.7	1.1	-	-	-
20/15------------------	33.9	55.6	53.5	49.0	23.0	5.1	1.2	-
20/20------------------	22.4	22.1	23.1	29.0	26.1	19.6	8.1	1.2
20/30------------------	15.5	9.4	7.4	8.8	22.9	27.2	23.4	17.6
20/40------------------	6.9	2.8	2.2	3.0	8.5	11.7	16.1	25.0
20/50------------------	4.6	1.3	1.5	1.9	5.7	9.4	10.7	11.1
20/70------------------	3.2	1.2	1.7	1.6	3.3	6.3	7.0	5.8
20/100-----------------	7.8	2.6	3.7	2.8	5.8	14.8	23.0	30.1
20/200-----------------	2.8	2.2	2.4	1.4	2.8	4.0	6.5	2.2
Less than 20/200--------	1.5	1.0	1.3	0.8	0.8	1.9	4.0	7.0
Women								
Total-------------	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
20/10 or better---------	0.8	1.8	1.1	1.3	0.4	-	-	-
20/15------------------	24.9	42.0	43.2	36.2	12.3	3.3	0.7	-
20/20------------------	24.8	26.9	31.0	36.5	26.7	14.2	2.2	1.6
20/30------------------	15.2	11.3	11.6	12.8	24.2	17.0	16.1	8.5
20/40------------------	6.2	4.4	2.7	3.3	7.2	12.0	11.1	10.3
20/50------------------	4.6	2.4	1.4	1.9	6.9	7.3	9.4	13.4
20/70------------------	3.8	1.0	1.4	1.4	4.3	7.8	6.8	21.2
20/100-----------------	11.2	4.4	2.0	2.8	11.1	25.1	33.3	26.1
20/200-----------------	5.3	3.6	2.3	2.3	4.7	11.1	11.4	9.8
Less than 20/200--------	3.2	2.2	3.3	1.5	2.2	2.2	9.0	9.1

Table 8. Percentage of Adults Reaching Specified Acuity Levels-Corrected Far Acuity.

Sex and acuity level	Total, 18-79 years	18-24 years	25-34 years	35-44 years	45-54 years	55-64 years	65-74 years	75-79 years
Both sexes	Percent distribution							
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
20/10 or better	1.5	2.4	2.6	1.7	0.9	0.5	0.2	-
20/15	38.4	57.8	57.7	52.8	31.1	11.2	4.4	1.6
20/20	32.9	27.5	29.3	35.5	39.0	38.4	27.6	13.4
20/30	17.7	10.2	7.4	7.1	20.5	33.5	38.9	38.6
20/40	4.5	1.4	1.5	1.4	4.2	8.5	12.6	18.0
20/50	1.8	0.4	0.3	0.8	1.5	2.6	6.2	9.3
20/70	0.9	-	0.4	0.3	0.5	1.6	2.0	7.9
20/100	1.5	0.0	0.5	0.2	1.6	2.5	5.3	7.9
20/200	0.4	0.2	0.2	0.1	0.4	0.7	0.9	1.3
Less than 20/200	0.4	0.1	0.1	0.1	0.3	0.5	1.9	2.0
Men								
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
20/10 or better	1.9	2.9	4.0	2.0	1.4	0.2	-	-
20/15	42.8	61.3	63.5	58.4	37.3	12.2	4.5	3.2
20/20	31.0	25.4	25.0	30.9	36.6	39.1	32.4	15.7
20/30	15.8	8.6	5.4	5.8	18.0	32.5	35.0	38.0
20/40	4.2	1.6	1.2	1.4	3.2	8.8	11.9	19.8
20/50	1.7	0.2	0.2	0.5	1.9	2.3	5.8	10.3
20/70	0.7	-	0.1	0.5	0.3	1.4	2.8	3.2
20/100	1.4	-	0.4	0.1	0.9	2.8	5.9	8.3
20/200	0.3	-	0.2	0.2	0.4	0.2	1.2	-
Less than 20/200	0.2	-	-	0.2	-	0.5	0.5	1.5
Women								
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
20/10 or better	1.1	1.9	1.3	1.5	0.5	0.8	0.4	-
20/15	34.7	54.9	52.6	47.5	25.3	10.3	4.3	-
20/20	34.6	29.2	33.2	39.8	41.1	37.8	23.8	11.2
20/30	19.5	11.6	9.3	8.2	22.9	34.4	42.1	39.1
20/40	4.7	1.2	1.7	1.5	5.2	8.3	13.1	16.4
20/50	1.8	0.6	0.3	1.0	1.1	2.8	6.4	8.4
20/70	1.0	-	0.7	0.1	0.7	1.8	1.4	12.5
20/100	1.6	0.1	0.6	0.3	2.3	2.2	4.8	7.4
20/200	0.4	0.3	0.2	-	0.3	1.2	0.7	2.6
Less than 20/200	0.6	0.2	0.1	0.1	0.6	0.4	3.0	2.4

Shown in Figure 39 is the distribution of binocular visual acuity scores. (Data on near acuity was not obtained.) Notice that approximately 90% of those tested had 6/6 acuity or better corrected normal vision, which is surprisingly high. Consistent with this, when Davison and Irving compared their results with other driver surveys, they found their results reported a smaller fraction of people with poor acuity.

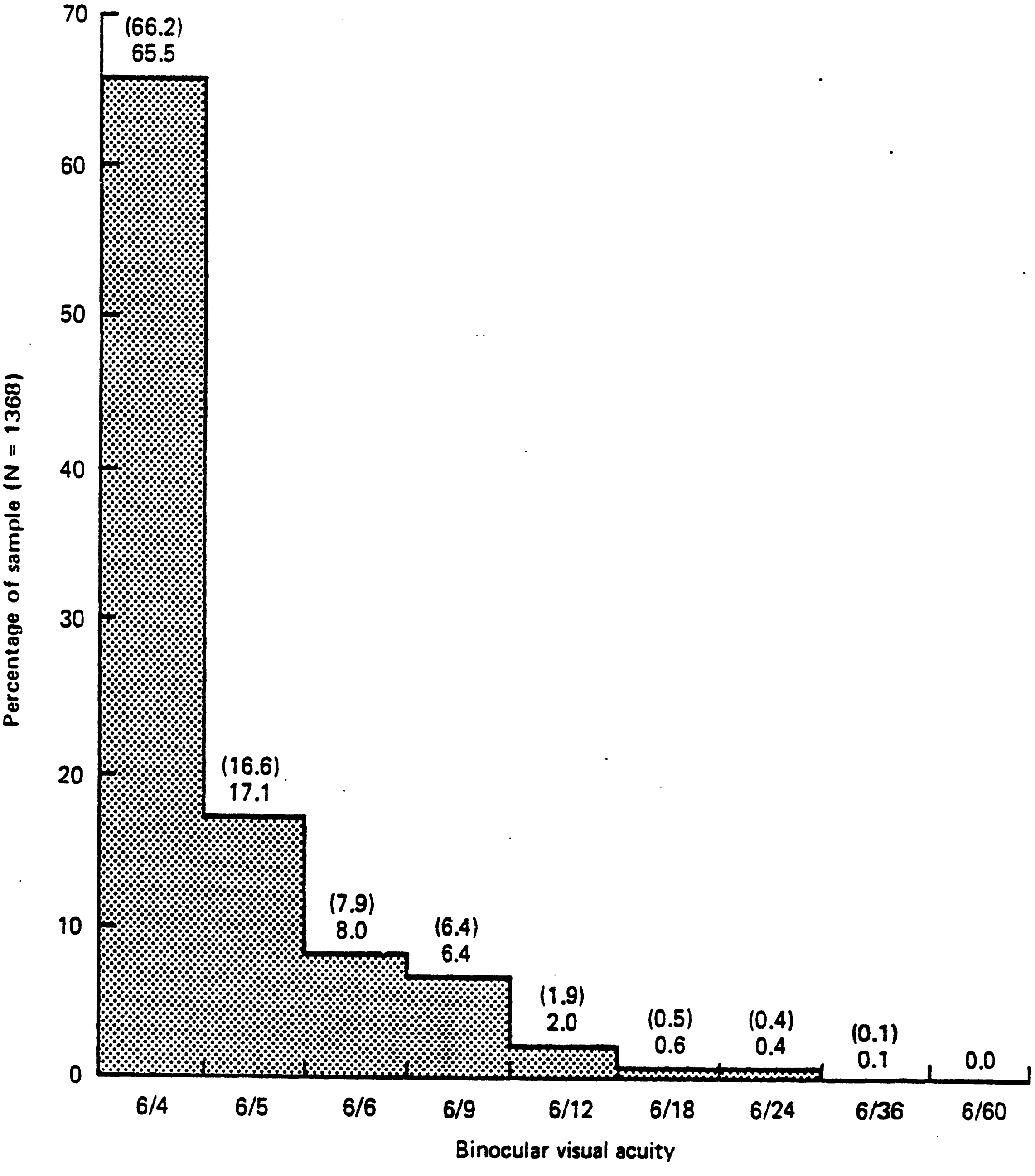


Figure 39. Distribution of Visual Acuity Scores from Davison and Irving (1980).

Figure 40 shows the mean acuity for different age groups and the fraction that fail to meet various acuity standards. The figure shows a clear degradation in acuity at about age 45. An implication of this is that the design specification for character sizes for instrument panel labeling should vary with the market segment for which a vehicle is intended. Using the adult averages for a product intended for older drivers is inappropriate.

Comparing the failure rate line (standard 6/9) in Figure 40 with subtotals from Table 8 (those values corresponding to 20/40 vision or worse for each age), may suggest that people in the U.S. (Table 5) do not see as well as those in the U.K. For example, the failure rate at age 75-79 is about 17% (Figure 40). Based on Table 5, the comparable figure is 36% (18 + 9.3 + ... + 2.0). One explanation is that the U.S. data is from all adults, whereas the British data is only from those adults who see well enough to drive. Further, the two surveys used different procedures for measuring acuity. Hence, data on adult visual acuity can be misleading when applied to drivers, especially for older populations or at extremes.

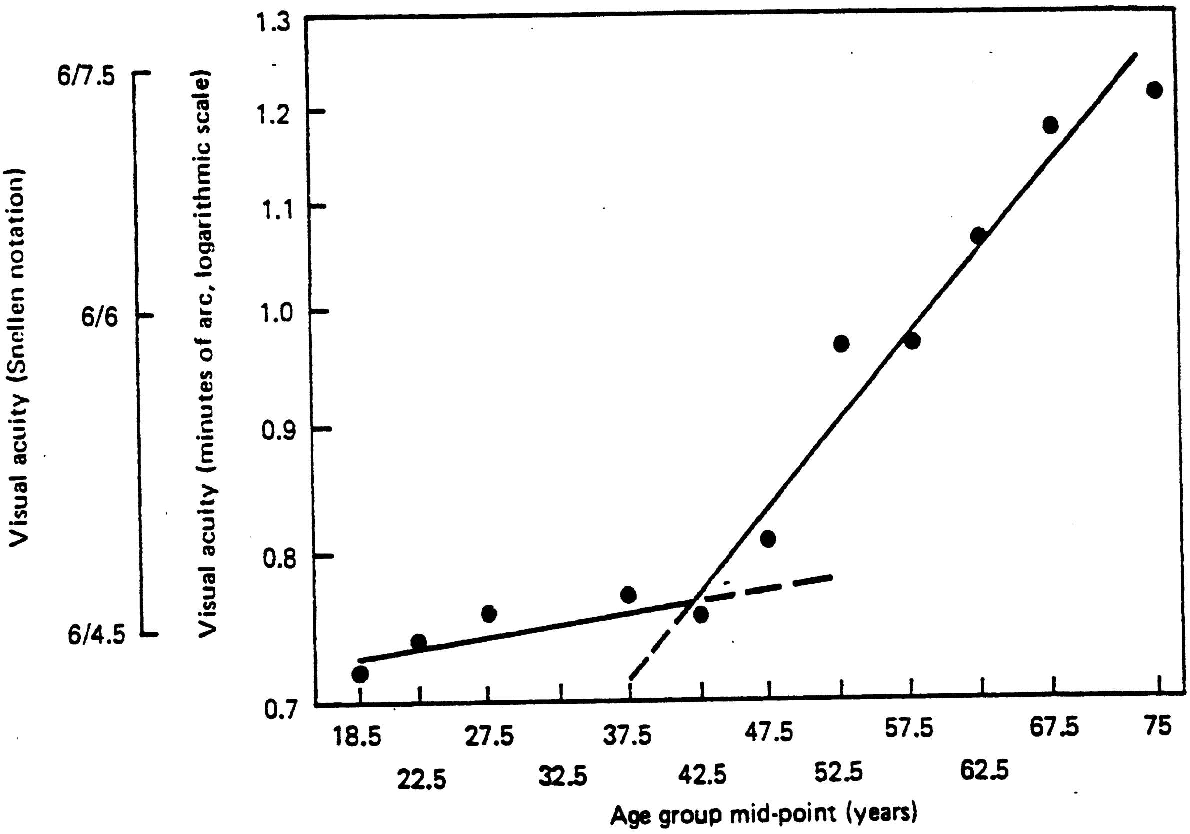


Figure 40. Mean Acuity As a Function of Age (Davison and Irving, 1980).

Summary

There are two very good studies in the literature of human visual acuity, one from the U.S., one from the U.K. They both provide detailed data on visual acuity as a function of age and show marked loss of acuity at age 45. However, they do not offer recommendations as to what level of acuity should serve as a design goal.

APPLICATION INDEPENDENT PREDICTIONS OF LEGIBILITY

Good research not only answers the specific questions which prompted the research, but also provides insights into answers for questions that were not asked. Further, good research not only identifies which of several alternatives is best and by how much, but why as well. This section presents examples of research papers which contain legibility predictions useful for a wide range of applications and conditions.

Peters and Adams (1959)

One of the most commonly referred to recommendations for character height is that of Peters and Adams (1959). That recommendation is as follows:

Letter Height (inches) = $H = .0022D + K1 + K2$

where:

D = Viewing Distance (inches)
K1 = Correction factor for illumination and reading situation
= 0.06 for illumination $>$ 1.0 fc, favorable reading conditions
= 0.16 for illumination $>$ 1.0 fc, unfavorable conditions or illumination $<$ 1.0 fc, favorable conditions
= 0.26 for illumination $<$ 1.0 fc, unfavorable conditions
K2 = Correction for Importance
= 0.075 for emergency labels, counters, scales, legend lights
= 0.0 for other (unimportant) panel markings

Peters and Adams note that for standard consoles, the viewing distance is 28 inches and K1 is 0.16 . That leads to recommended heights of just over .22 inches for unimportant markings and almost .30 inches for important ones.

These recommendations are often cited because they appeared in a popular edition of the de facto standard human factors textbook (McCormick, 1970, pages 170-171). In spite of their wide use, it is not clear if these recommendations are supported by empiric data. The Peter and Adams paper does not cite any original research though it is believed (from reading Van Cott and Kinkade, 1972) that these recommendations emerged from the work of Brown and others described elsewhere in this report.

Duncan and Konz (1976)

In Duncan and Konz (1976) (see also Duncan, 1977), eight men viewed two 7-segment and one 16-segment light-emitting diode (LED) displays, and two 7-segment liquid-crystal displays (LCD). The five displays varied in many ways (font, height, height to strokewidth ratio, etc.). Figure 41 shows examples of the displays. The five different displays were mounted in a grey panel which was flush-mounted at eye-height on a light green wall.

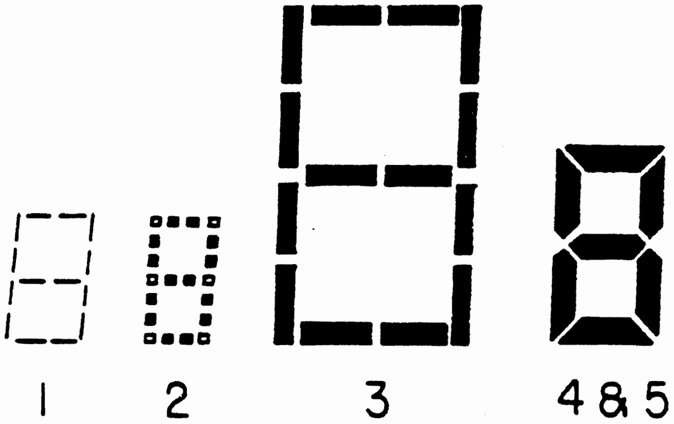


Figure 41. Displays Examined by Duncan and Konz (1976).

In the first condition, the visual angle was held constant at 31 minutes of arc. Digits were shown one at a time and participants said which was shown when they were "100% confident." Each of the 10 digits was displayed three times at each of three illumination levels (16, 151, and 484 lux (1.5, 14, 45 fc)) for each of the five displays.

Response times for the LED displays (543 ms) were significantly less than those for the LCDs (599 ms). For the LCDs, response times for the reflective displays (557 ms) were less than those for the transmissive LCDs (640 ms) for the conditions examined. This does not imply that one technology is superior to another, because the displays varied in many ways so that they were not directly comparable. Further, the LCDs had rise times on the order of 50 msec and delay times between 70 and 150 msec. They were almost zero for the LEDs, making the true differences in human response to these two displays of no practical significance.

In the second condition, people adjusted the illumination up from 0 and down from 1291 lux (120 fc) to a "preferred" ambient level. For the LEDs, preferred illumination levels varied between 108 and 226 lux (10 and 21 fc). For the reflective and transmissive LCDs, the preferred values were 667 and 291 lux (62 and 27 fc), respectively.

In the third condition, people started from 15.5 meters away and approached each display until 5 two-digit numbers were read without error. The same three illumination levels used in the first condition were examined here. Displays were read without error at visual angles ranging from 4.8 to approximately 9 minutes of arc with the viewing angle tending to be less for LEDs. (See Figure 42.)

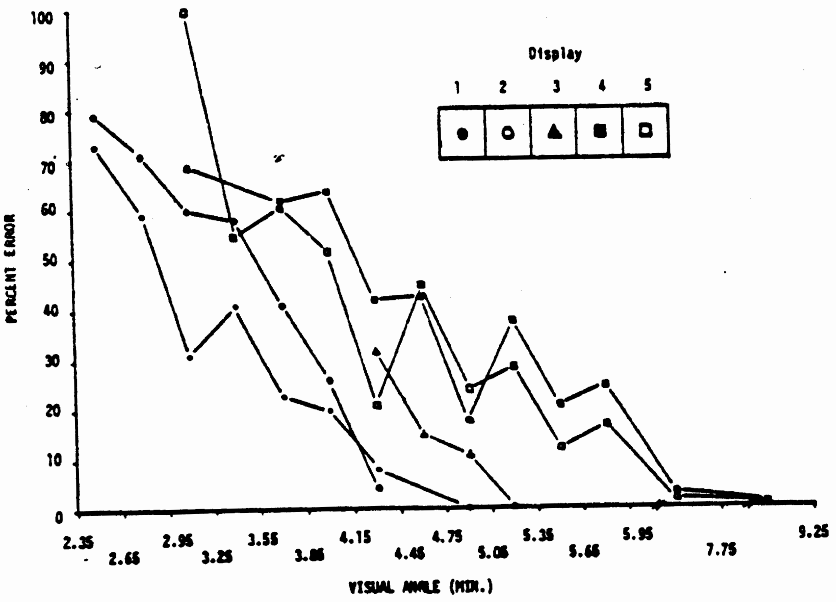


Figure 42. Percent Errors in Duncan and Konz (1977).

Based on a regression analysis of the data, the no-error viewing distance in meters can be found using:

No-error Viewing Distance (cm) =
De = 243.2 + 695H - 36.1*(H:Sw)

where:

H = Character Height (centimeters)
H:SW = Height:Strokewidth Ratio.

It should be noted that the above equation has been modified from those appearing in the technical report. First, Duncan and Konz use strokewidth to height ratio rather than the inverse used here. Second, the Duncan and Konz formula gave viewing distance in <u>meters</u> as a function of character height in <u>millimeters</u>, where the equations above are expressed in consistent units (centimeters). Further, these equations accounted for 72 and 83% of the variance, respectively. None of the other factors (font, luminance contrast, illumination, display luminance) led to a significant reduction in the variance in a stepwise regression analysis.

For the purposes here, a version of the equation with height as the predicted variable is more useful. The re-arranged equation appears below:

$$H = .0015De + .0519(H:Sw) - .3499$$

It should be noted that the errors were not randomly distributed. Tables 9 and 10 show the number of errors for each digit and viewing angle. Table 11 identifies the confusions that were common. The digits 0, 8, and 9 had high error counts and were commonly confused with each other. This outcome agrees with the results of Van Nes and Bouma described previously.

Table 9. Percent Error When Reading LED Displays in Duncan (1977).

VISUAL ANGLE RANGE (MIN.)	(SAMPLE SIZE FOR DISPLAY-1) n_1	DIGITS OF DISPLAY-1										(SAMPLE SIZE FOR DISPLAY-2) n_2	DIGITS OF DISPLAY-2										(SAMPLE SIZE FOR DISPLAY-3) n_3	DIGITS OF DISPLAY-3									
		0	1	2	3	4	5	6	7	8	9		0	1	2	3	4	5	6	7	8	9		0	1	2	3	4	5	6	7	8	9
2.35-2.65	15	47	13	33	67	7	47	33	47	73	73	19	89	16	32	53	32	68	58	26	79	84	--	--	--	--	--	--	--	--	--	--	--
2.65-2.95	27	26	0	22	37	0	22	22	19	59	48	24	79	8	12	25	12	50	42	8	71	67	--	--	--	--	--	--	--	--	--	--	--
2.95-3.25	13	23	0	8	15	0	15	38	15	31	15	10	30	0	0	10	10	10	30	10	60	50	--	--	--	--	--	--	--	--	--	--	--
3.25-3.55	29	17	7	7	17	0	14	31	21	41	28	31	55	3	3	16	16	16	16	16	58	29	--	--	--	--	--	--	--	--	--	--	--
3.55-3.85	22	5	0	5	9	0	5	0	0	23	32	22	32	0	0	9	5	14	18	9	41	23	--	--	--	--	--	--	--	--	--	--	--
3.85-4.15	20	0	0	0	10	0	5	5	5	20	10	19	11	0	0	5	0	0	5	5	26	5	--	--	--	--	--	--	--	--	--	--	--
4.15-4.45	24	0	0	0	8	0	4	4	0	8	0	24	8	4	0	0	0	0	0	4	4	0	25	0	12	8	24	4	8	24	20	32	36
4.45-4.75	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	34	3	6	6	15	3	6	21	32	15	12
4.75-5.05	24	0	0	0	0	0	0	4	0	0	0	--	--	--	--	--	--	--	--	--	--	--	28	0	7	4	11	0	0	11	18	11	11
5.05-5.35	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	18	0	6	6	6	0	6	11	22	0	6
5.35-5.65	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
5.65-5.95	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
5.95-7.75	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
7.75-9.25	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 10. Percent Error When Reading LCD Displays in Duncan (1977).

VISUAL ANGLE RANGE (MIN.)	(SAMPLE SIZE FOR DISPLAY-4) n_4	DIGITS OF DISPLAY-4										(SAMPLE SIZE FOR DISPLAY-5) n_5	DIGITS OF DISPLAY-5									
		0	1	2	3	4	5	6	7	8	9		0	1	2	3	4	5	6	7	8	9
2.35 - 2.65	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2.65 - 2.95	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2.95 - 3.25	13	77	8	31	46	31	54	23	77	69	92	8	88	62	62	88	25	62	50	62	100	100
3.25 - 3.55	17	65	12	29	41	18	29	29	59	65	100	11	82	27	36	55	18	64	45	45	55	91
3.55 - 3.85	26	58	8	42	38	12	15	27	35	62	85	23	70	39	17	65	17	43	43	52	61	87
3.85 - 4.15	28	36	0	14	21	4	7	29	50	64	71	29	59	41	21	34	28	38	41	41	52	86
4.15 - 4.45	12	0	0	0	0	0	0	0	8	42	50	14	29	21	7	29	14	14	21	43	21	71
4.45 - 4.75	21	29	0	10	14	10	10	14	19	43	43	22	50	23	41	45	27	27	36	41	45	64
4.75 - 5.05	17	6	0	6	0	0	0	6	0	18	12	17	18	18	0	29	18	18	24	35	24	59
5.05 - 5.35	21	14	0	0	0	5	10	5	14	38	24	21	29	19	19	24	19	19	24	29	29	48
5.35 - 5.65	19	0	0	0	0	0	0	0	0	21	5	17	6	12	0	18	6	6	6	18	12	29
5.65 - 5.95	24	4	0	4	4	0	0	8	4	25	21	23	17	4	4	26	9	9	17	22	17	30
5.95 - 7.75	32	0	0	0	0	0	0	0	0	19	22	28	4	14	4	7	4	4	4	11	11	14
7.75 - 9.25	25	0	0	0	0	0	0	4	0	4	0	-	-	-	-	-	-	-	-	-	-	-

Table 11. Digit Confusions Reported by Duncan (1977).

DIGIT DISPLAYED	DISPLAY-1		DISPLAY-2		DISPLAY-3		DISPLAY-4		DISPLAY-5	
	VAR (min)	Digits Read	VAR (min)	Digits Read	VAR (min)	Digits Read	VAR (min)	Digits Read	VAR (min)	Digits Read
0	3.25-3.85	8*,6,9,4	3.85-4.45	8*,4,6	4.15-4.75	1*	5.35-5.95	5*	5.95-7.75	3*
1	2.95-3.55	4*	3.85-4.45	4*	4.75-5.35	0*,7	3.25-3.85	4*,7,3	5.95-7.75	6*,2,0
2	3.25-3.85	4*,3	2.95-3.55	4*	4.75-5.35	5*,7	5.35-5.95	7*	5.95-7.75	4*
3	3.85-4.45	4*,9,1	3.55-4.15	7*,2,6	4.75-5.35	7*,9,6	5.35-5.95	5*	5.95-7.75	4*,6
4	2.35-2.65	2*	3.25-3.85	7*,1,5,3	4.15-4.75	1*,9	4.75-5.35	2*	5.95-7.75	1*
5	3.85-4.45	9*,6	3.25-3.85	6*,8,9,2	4.75-5.35	3*	4.75-5.35	2*,4	5.95-7.75	6*
6	4.45-5.05	5*	3.55-4.15	4*,2,1	4.75-5.35	7*,0,3,1	7.75-9.25	7*	5.95-7.75	5*
7	3.55-4.15	1*	3.85-4.45	3*,1	4.75-5.35	1*,3,0	5.35-5.95	1*	5.95-7.75	1*,2
8	3.85-4.45	0*,9,4,2	3.85-4.45	6*,0,9	4.45-5.05	3*,2,6, 4,0	7.75-9.25	0*	5.95-7.75	3*,2
9	3.55-4.15	5*,4,3	3.55-4.15	5*,3,4,7,2	4.75-5.35	4*,5	5.95-7.75	4*,5,7	5.95-7.75	4*,7

*Predominant misreading

In the fourth condition, people walked to the location where they would prefer to view the display if they were required to do so once per minute for an hour. Again, the three illumination levels used in condition one were examined. Using regression analysis, the following predictions for preferred viewing distance were computed:

Preferred viewing distance (cm) =
$$Dp = 22.9 + 265(H) - 10.2(H:SW)$$

This equation accounts for 73% of the variance. The ratio of preferred to no-error viewing angles is about 4.7:1. Duncan and Konz refer to another study (Fortuin, 1970) that suggests that objects should be more than 2.5 times the minimum detectable size for "easy seeing." In regard to preferences, participants liked the LED displays the most except at the high illumination condition, where there were no differences.

As before, the equation above can be re-arranged to give predictions for required character height.

$$H = .0038Dp + .0385(H:Sw) - .0864$$

A comparison of predicted letter heights between the no error viewing distance and preferred viewing distance equations and several other methods can be found later in this section.

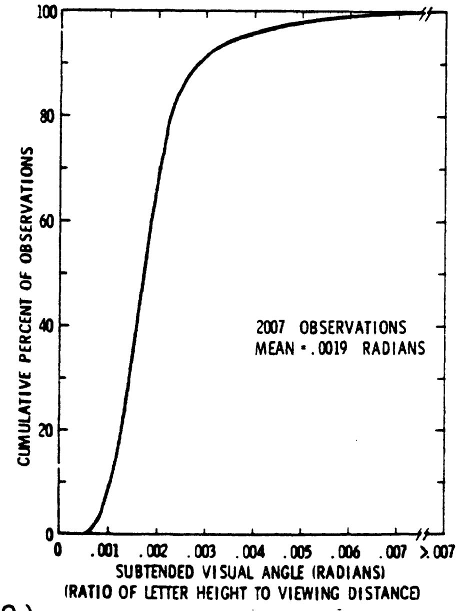
The authors conclude that for ease of viewing and maximum legibility a 20-minute visual angle is recommended for LED displays. For LCDs, 26 minutes of arc is optimal for transmissive displays, while 30 minutes of arc is optimal for reflective LCDs.

Smith (1979)

Smith reports data from 88 student experimenters who took a class from him. They tested 547 viewers to find the maximum reading distance legibility of 314 different sample test materials. Viewers walked up to the materials and said when they could read them. This procedure of finding maximum reading distance is identical to many early studies of highway sign legibility (Forbes and Moscowitz, 1950, Kuntz and Sleight, 1950, and Allen, Smith, Janson, and Dyer, 1966) as described in Zeltner, Ratanaproeksa, Goldstein, Adams,and Green (1988). Many different display materials, including newsprint, magazine advertisements, and company letterheads, were used. These materials encompassed a wide variety of fonts, stroke widths, and spacing. The people tested covered a wide range of visual acuities and ages, and the testing occurred under a variety of viewing conditions.

From the distribution of 2007 responses, it can be seen in Figure 43 that all but eight of the test materials were legible when the subtended viewing angle was .007 radians or less. Using this result, and the fact that for angles less than seven degrees the sine, tangent, and angle measure in <u>radians</u> are all equal (to three significant figures), the following formula to predict character height was developed:

Height = .007 x Viewing Distance (D, same units as height)



Source: Smith (1979)

Figure 43. Cumulative Distribution of Visual Angle at Limit of Legibility.

Because of the value of the constant (007), this relationship is known as the James Bond Rule.

Howett (1983)

Howett derived a procedure to compute the minimum letter height for legibility given the viewing distance and the observer's visual acuity. The mathematical derivation assumes that legibility depends only on the visual angle of the object, its luminance, and its luminance contrast. This is generally true for distances in excess of one meter. The expression is based upon acuity data reported by Nakane and Ito (1978) and Kaneko (1982). Readers interested in the details of the derivation should read Howett's technical report.

The procedure to determine letter height is as follows. First, compute the luminance contrast.

$$\text{Contrast } (\%) = C = ((Lb - Lt) / Lb) \times 100 \quad \text{(assumes } Lb>Lt\text{)}$$

where: Lb = Background Luminance
Lt = Target Luminance

Since C is dimensionless, the units of Lb and Lt do not matter here as long as they are the same.

Second, compute the relative Snellen Acuity, S. The Snellen acuity is a measure of how well people see. It was described earlier. In this instance, the acuity of interest is that of the worst case viewer.

$$S = Sd * (85 / Lb)^{.213} * (90 / C)^{.532}$$

where: Sd = Denominator in the Snellen ratio.
(If a viewer has 20/40 visual acuity, use 40.)
Lb = Background Luminance (cd/m2)

The third step is to compute the character height, H.

$$H = (H:Sw) * 1.45 * 10^{-5} * S * D$$

where: H:Sw = Height to Strokewidth Ratio (for 6:1 use 6)
D = Viewing Distance (m)

Payne (1983)

This study investigated the effect of viewing angle, level of back light, character subtense, and ambient light level on reading error rates of four-digit, seven-segment, reflective liquid crystal displays. All four factors were viewed by 120 people at five different levels for 0.5 seconds. Participants called out the number displayed trying "to be as accurate as possible." Viewing angle varied uniformly from 0 to 60 degrees, back light luminance varied uniformly from 0 to 122 cd/m^2 (0 to 35.6 fL), character subtense varied from 0.25 to 1.43 degrees, and ambient light levels varied from 20 to 1500 lux (1.9 to 140 fc). A total of 24,000 data points were collected.

A multiple regression analysis revealed error rates could be estimated as follows:

$$\text{Error Rate (\%)} = E = 1.52 + .02Bl - 1.40Ca + .02Va - .0006Ea$$

Where:

Bl = Back Light Luminance (0 to 122 cd/m^2)
Ca = Character Subtense Angle (0.025 to 1.34 degrees)
Va = Viewing Angle (0 to 60 degrees)
Ea = Ambient Light Illumination (20 to 1500 lx).

As can be seen from this equation, error percentage rates increase as back light and viewing angle increase, and decrease as character subtense (size) and ambient light decrease. However, the most significant coefficient in this equation is character size, which is 70 times greater than any of the other coefficients in this equation.

It is noted by the author that the regression equation is more useful in making comparisons of different display situations rather than being used as an absolute predictor of error rate.

Sawyer and Talley (1987)

Sawyer and Talley (1987) report recommendations for character height which have become popular with designers and have been widely distributed by the authors. Their recommendations are summarized in Table 12. Their recommendations were based upon their interpretation of "an amalgamation of the literature" and assume that the contrast ratio is high (Talley, 1988). Their numbers seem reasonable. For example, for a 28-inch viewing distance, a typical value for automobile instrument panels, the tallest character in the recommended range is .26 inches (29.8 minutes of arc). This is close to the value from the Bond Rule (.007 radians, 24 minutes of arc). At the other end of the range, 17.2 minutes (.14 inches) is identified as the smallest size preferred and 12.3 minutes (.10 inches) is the minimum recommended height.

Table 12. Character heights recommended by Sawyer and Talley (1987).

Viewing Distance in inches (centimeters)

Character Height in. (cm)	18.0 (45.7)	20.0 (50.8)	22.0 (55.9)	24.0 (61.0)	26.0 (66.0)	28.0 (71.1)	30.0 (76.2)	32.0 (81.3)	34.0 (86.4)	36.0 (91.4)
0.04 (0.10)	7.6	6.9	6.3	5.7	5.3	4.9	4.6	4.3	4.0	3.8
0.06 (0.15)	**11.5**	**10.3**	**9.4**	8.6	7.9	7.4	6.9	6.4	6.1	5.7
0.08 (0.20)	**15.3**	**13.8**	**12.5**	**11.5**	**10.6**	9.8	9.2	8.6	8.1	7.6
0.10 (0.25)	**19.1**	**17.2**	**15.6**	**14.3**	**13.2**	**12.3**	**11.5**	**10.7**	**10.1**	9.6
0.12 (0.30)	**22.9**	**20.6**	**18.8**	**17.2**	**15.9**	**14.7**	**13.8**	**12.9**	**12.1**	**11.5**
0.14 (0.36)	**26.7**	**24.1**	**21.9**	**20.1**	**18.5**	**17.2**	**16.0**	**15.0**	**14.2**	**13.4**
0.16 (0.41)	30.6	**27.5**	**25.0**	**22.9**	**21.2**	**19.6**	**18.3**	**17.2**	**16.2**	**15.3**
0.18 (0.46)	34.4	30.9	**28.1**	**25.8**	**23.8**	**22.1**	**20.6**	**19.3**	**18.2**	**17.2**
0.20 (0.51)	38.2	34.4	31.3	**28.7**	**26.4**	**24.6**	**22.9**	**21.5**	**20.2**	**19.1**
0.22 (0.56)	42.0	37.8	34.4	31.5	**29.1**	**27.0**	**25.2**	**23.6**	**22.2**	**21.0**
0.24 (0.61)	45.8	41.3	37.5	34.4	31.7	**29.5**	**27.5**	**25.8**	**24.3**	**22.9**
0.26 (0.66)	49.7	44.7	40.6	37.2	34.4	31.9	**29.8**	**27.9**	**26.3**	**24.8**
0.28 (0.71)	53.5	48.1	43.8	40.1	37.0	34.4	32.1	30.1	**28.3**	**26.7**
0.30 (0.76)	57.3	51.6	46.9	43.0	39.7	36.8	34.4	32.2	30.3	**28.7**
0.32 (0.81)	61.1	55.0	50.0	45.8	42.3	39.3	36.7	34.4	32.4	30.6

Visual Angle in minutes of arc

Underlined area is the preferred range of visual angle.
Tinted area is the minimum recommended range.

Summary

Two methods are presented in this section are for determining required character height for virtually error-free reading from many types of characters, two methods for which accuracy is unknown, and one method for predicting error rates in reading seven segment numerals. In all cases, it is either shown or assumed that the primary factor affecting legibility is the visual angle of the character. While viewing distance does matter, its effect is secondary.

Shown in Table 13 is a comparison of predicted letter heights for a variety of conditions using six methods described in this report. The data in this Table were plotted and are shown in Figure 44. The prediction equations used from Peters and Adams (1959), Duncan and Konz (1976), Smith (1979), and Howett (1983) appear earlier in this section. A prediction methodology proposed by Farber (1988) appears in the Blackwell's Research section, while letter height predictions for broken circle test objects proposed by Shlaer, Smith, and Chase (1942) appears in the Basic Legibility Relationships section. A seventh method, derived from Mourant and Langolf (1976) can be used to predict letter height as a function of task reaction time. This method was not included in Table 13 because the other methods included are based on threshold measurements, not reaction times. Consequently, comparing predicted letter heights using the Mourant and Langolf method would be purely subjective based on the reaction time chosen.

Table 13. Comparison of Predicted Required Letter Heights Needed for Legibility.

	Predicted Letter Height (cm)					
Condition[a]	Farber	Smith	Howett	Shlaer	Peters & Adams	Duncan & Konz[b]
1	0.39	0.70	0.35	0.74	0.56	0.11, 0.52
2	3.90	7.00	3.48	7.36	2.54	1.40, 3.92
3	38.6	70.0	34.8	73.4	22.3	14.4, 37.9
4	0.40	0.70	0.39	0.75	0.56	0.11, 0.52
5	4.00	7.00	3.89	7.50	2.54	1.40, 3.92
6	40.1	70.0	38.9	75.0	22.3	14.4, 37.9
7	0.60	0.70	0.53	0.75	0.56	0.11, 0.52
8	6.00	7.00	5.33	7.50	2.54	1.40, 3.92
9	60.2	70.0	53.3	75.0	22.3	14.4, 37.9

[a] Condition 1 assumes very black characters on a very white background (contrast assumed to be 90%), sign luminance = 85 cd/m^2 (25 fL), maximum viewing distance is 1 meter, target population visual acuity of 20/40 (the legal driving visual acuity in many states), a 6:1 height to strokewidth ratio (a typical ratio found in many display situations), and characters which convey "important information."

Conditions 2 and 3 are the same as condition 1, except that maximum viewing distances are 10 and 100 meters, respectively.

Conditions 4 - 6 are the same as condition 1, except that sign luminance is 50 cd/m^2 (14.6 fL), and the maximum viewing distance is 1, 10, and 100 meters, respectively.

Conditions 7 - 9 are the same as conditions 4 - 6, except that the display does not have very black letters on a very white background, contrast is 50%, and the maximum viewing distance is 1, 10, and 100 meters, respectively.

[b] For a given condition, the first number is the letter height required for error-free performance, while the second letter height is based on test participant subjective preference data.

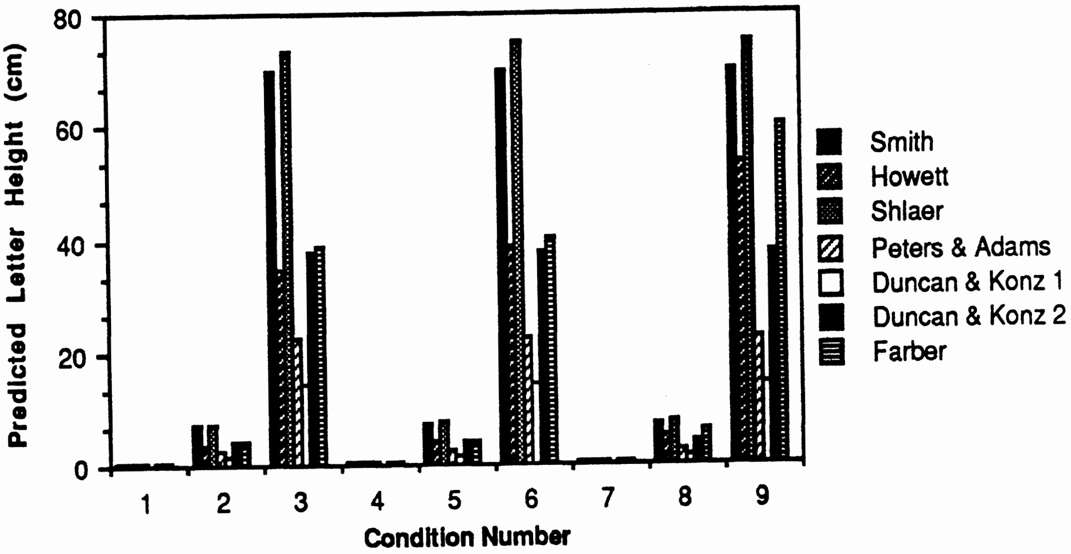


Figure 44. Predicted Letter Height for a Variety of Methods and Conditions.

As can be seen from Table 13 and Figure 44, the Smith and Shlaer et al. methods predict virtually the same character height throughout the range of conditions examined, but vary considerably from letter heights predicted using the other four methods. The letter height predictions using the Howett and Farber methods are almost identical across all conditions. Further, the Howett and Farber predictions are typically about half the size of the Smith and Shlaer et al. predictions throughout the entire range of distances tested. The Peters and

Adams formula consistently predicts letter heights between 25% smaller for the 1-meter viewing conditions and 236% smaller for the 100-meter viewing condition than those predicted using the Smith and Shlaer et al. methods. However, this method predicts slightly smaller letter heights than the Howett and Farber methods for viewing distances greater than 1 meter, while for the 1-meter distance, the Peters and Adams prediction is midrange between the Howett and Farber predictions and the Smith and Shlaer et al. predictions. It should be noted that the Peters and Adams predictions are based on constants whose correct usage is ambiguous at best, making the predictions somewhat suspect.

For the Duncan and Konz predictions, the required letter heights for error free performance is between 160% and 373% smaller than the letter heights which people would have subjectively preferred. This suggests that whenever feasible, letter heights should be much larger than those simply required for error-free performance, so that user preferences are met. In addition, the letter heights required for error-free performance using the Duncan and Konz method are many times smaller than letter heights predicted using any of the other methods. However, the Duncan and Konz preferred letter heights compare favorably with the letter heights predicted using the Howett, Farber, and the Peters and Adams formulas.

It should be emphasized that the predicted letter heights using the Duncan and Konz method for viewing distances of 100 meters (i.e., for conditions 3, 6, and 9) were outside the range of viewing distances used to develop both regression equations on which these predictions were made. Therefore, these letter height equations may or may not be valid for the 100-meter viewing distance.

The differences in predicted letter heights emphasizes that different prediction equations consider different factors, which in turn causes some variance between letter height prediction methods. Further, different definitions of minimum required letter height (i.e., letter height for error-free performance versus threshold legibility letter height) also accounts for some of the differences in predicted letter heights.

Further, as can be seen in Table 14 and Figure 44, the Shlaer et al. method produces reasonable predicted letter heights despite being first reported in 1942. This point again emphasizes that old research is not necessarily obsolete, but can still be used to provide useful insights and information.

It should also be noted that all conditions listed above were not directly taken into account in the development of every letter height prediction rule. That is, each rule was developed using only a subset of the conditions listed.

STUDIES OF HIGHWAY SIGNS

Due to inherent safety problems associated with driving, and the need to comprehend critical information quickly and accurately, highway sign legibility has been studied extensively. Clearly, being able to save lives and reduce accidents by minimizing driver errors caused by illegible signs makes this area of legibility research popular. Many of the findings resulting from studies of highway signs can be applied to other types of displays, such as those on automobile instrument panels.

The eight articles chosen for this section include some of the early, fundamental studies of highway sign legibility, as well as more recent, particularly well done and relevant studies.

Uhlaner (1941)

This study examined the strokewidth of three-inch block letters (height equals width) as a factor in legibility of highway signs. Dull black letters on a white background were tested outdoors under daylight conditions (2688 - 5914 lux, 250 - 550 fc).

In the first experiment 16 participants made a total of 1344 observations. Visual acuity was tested but not reported. Four representative letters (E, N, C, P) were tested using strokewidths which varied between 8% and 32% of letter height. Three observations were made of each letter-strokewidth combination, two while approaching, one while withdrawing. The results of the first study indicated that the optimal strokewidth is 18% of letter height. Legibility distance was a parabolic function of strokewidth:

$$\text{Legibility Distance} = 116 + 1236(Sw/H) - 3370(Sw/H)^2$$

where:

Sw = Strokewidth of the Character (inches)
H = Height of the Character (inches).

A second experiment was conducted to verify the results of the first. Fifteen participants made 1080 observations. Two groups of letters were examined (E, N, C, P and F, Z, B, O) using strokewidths of 16%, 18%, and 20%. Again, 18% of letter height was found to be the optimal strokewidth.

Smyth (1947)

This study attempted to find the ideal brightness of lettered signs when seen against different background brightness levels. A projected view of a street scene was used as the primary material in the experiment. A photograph of a "NO

THROUGH ROAD" sign was mounted in the screen, backed by a panel with variable brightness. Six people were tested, all with normal vision. The participants adjusted the sign letter brightness to correspond to three different criteria: minimum brightness at which the sign could be read; maximum brightness acceptable without discomfort or loss of background detail due to glare; and "ideal" brightness.

Results presented in Figure 45 showed that for the brightest scene (15 lux (1.4 fc)), maximum and ideal letter brightness levels were separated by a ratio of 10:1, whereas for the darkest scene (0.015 lux (0.0014 fc)) the ratio is only 4:1. The values for minimum brightness are 1/10 of the ideal letter brightness. The mean ideal letter brightness rises from 75 lux (7 fc) for the darkest scene to 269 lux (25 fc) for the brightest scene. There appears to be a fairly linear relationship between "ideal" letter brightness and the logarithm of the background brightness.

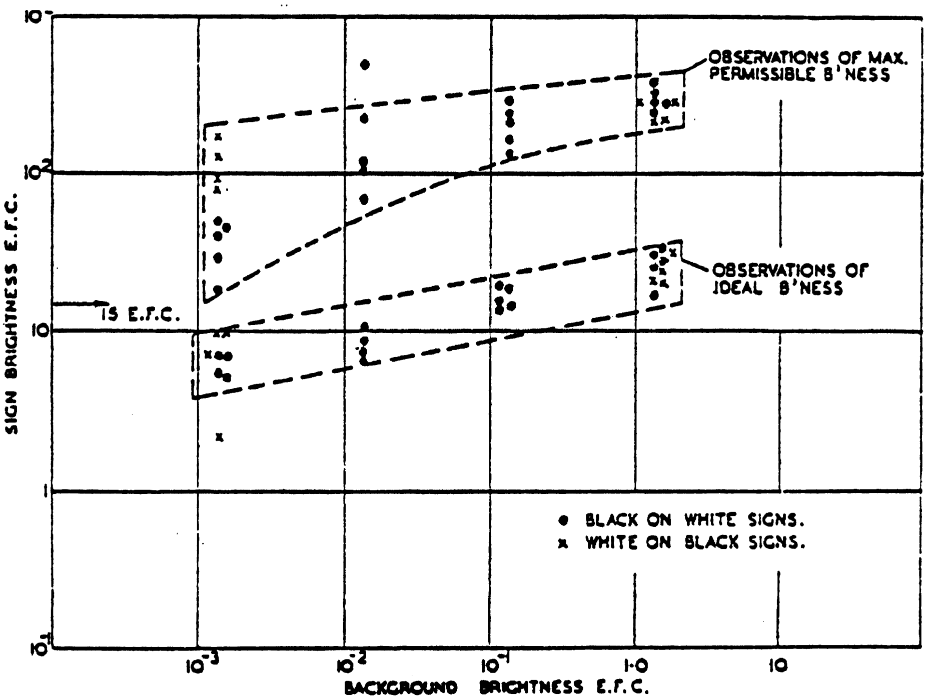


Figure 45. Relation Between Sign Brightness and Background Brightness.

The ideal brightness for legibility was found to be unaffected by the color of the sign, and was the same for black-on-white or white-on-black signs.

Hind, Tritt, and Hoffman (1976)

This study examined how the factors affecting sign legibility interact. The factors investigated were: font design, height to width ratio, height to strokewidth ratio, numeral/background contrast, luminance, grouping of numerals, and the distance between the display and observer. The various conditions investigated, based on previous research, were: six different numeral sets (the NAMEL, Berger, Motorways, Mackworth, Standards Association of Australia, and Tritt sets); height to strokewidth ratios of 15.8:1, 12:1, 10:1, 8:1, 6:1; contrast direction; and luminance levels of 1.0, 5.7, 34, 206, and 1233 cd/m^2 (0.29, 1.7, 9.9, 60, 360 fL).

The experiments were conducted in a dark tunnel of the Victorian Country Roads Board. Twenty-one people sat in four rows of chairs (5.5 to 11 m from a screen) and viewed slides of highway signs. Each slide was shown for 2.5 seconds. Each participant wrote down the numeral shown on the slide. There were a total of 252,000 responses.

The seven figures that follow illustrate the overall effects of the experimental variables and some of the major interactions. Figures 46 and 47 show the interaction between distance, luminance, height to strokewidth and contrast for both contrast directions. For the black-on-white contrast direction, increases in height to strokewidth ratio increased the percent recognized for all luminances tested. The opposite effect was found for the white-on-black contrast direction.

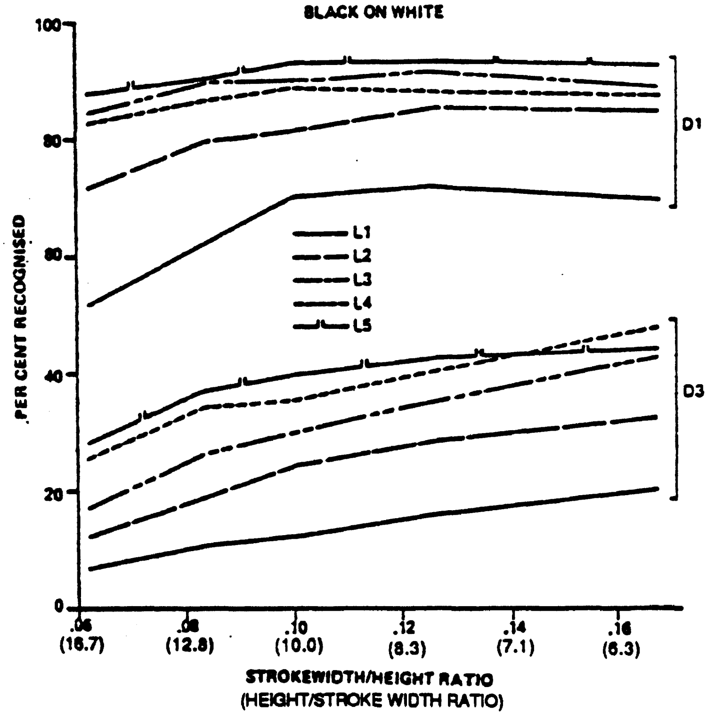


Figure 46. Illustration of the Effects of Distance, Luminance, and Strokewidth/height Ratio on Recognition of Black Numerals.

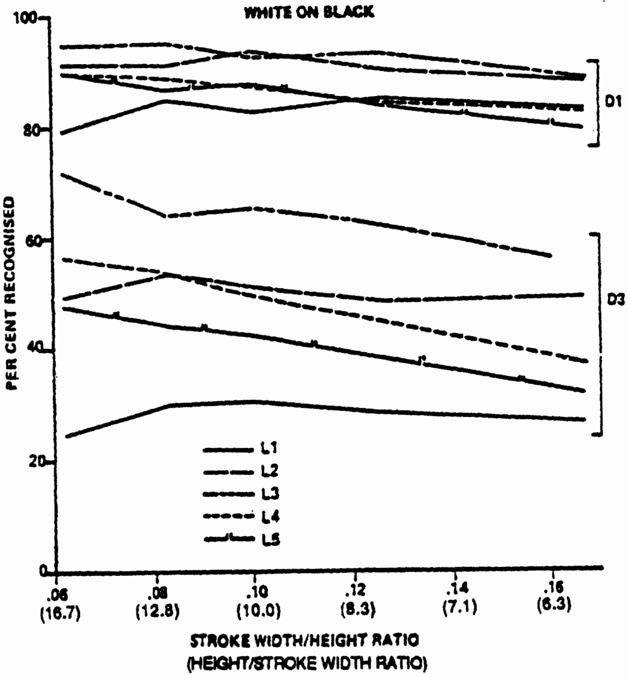


Figure 47. Illustration of the Effects of Distance, Luminance, and Strokewidth/height Ratio on Recognition of White Numerals.

Figure 48 shows that differences in luminance had a large effect on optimum height to strokewidth ratio for white characters, and virtually no effect on black characters. Further, it was found that for the range of luminances tested, the optimal height to strokewidth ratio for black characters ranged between two and three times greater than the optimal height to strokewidth ratio for white characters.

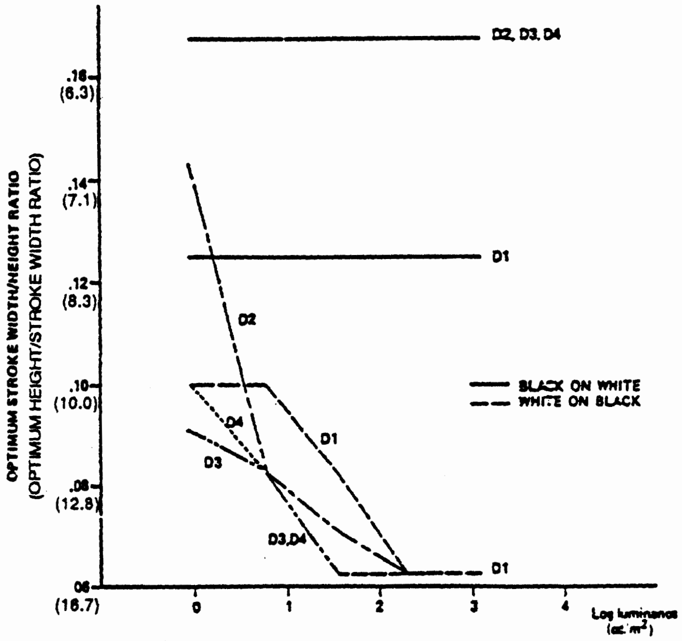


Figure 48. Optimum Numeral Strokewidth/height Ratios for Different Levels of Luminance and Distance.

Figures 49 and 50 show the interaction between font and height to strokewidth ratio. As can be seen from Figure 49, changes in fonts give absolute improvement in terms of percent recognition for white characters. This is not the case for black letters, as can be seen in Figure 50. Further, a decrease in height to strokewidth ratio decreases the percent recognition for white characters (Figure 48), but increases the percent recognition for black characters (Figure 49).

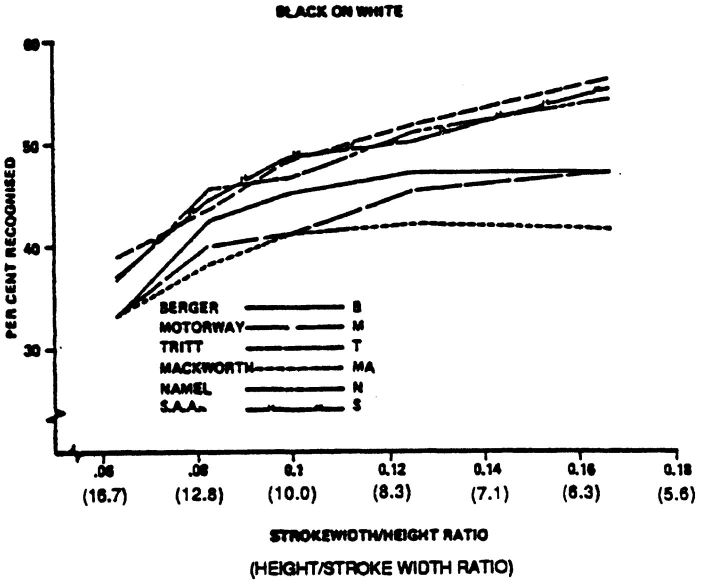


Figure 49. Effect of Font and Strokewidth/height Ratio on the Percentage of Black Numerals Recognized.

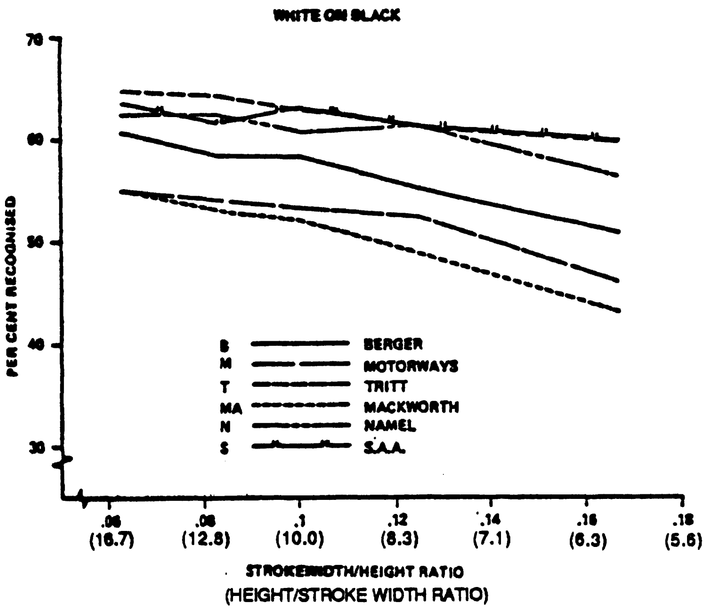


Figure 50. Effect of Font and Strokewidth/height Ratio on the Percentage of White Numerals Recognized.

Figures 51 and 52 show the interaction between luminance and distance. For black characters (Figure 51), increases in

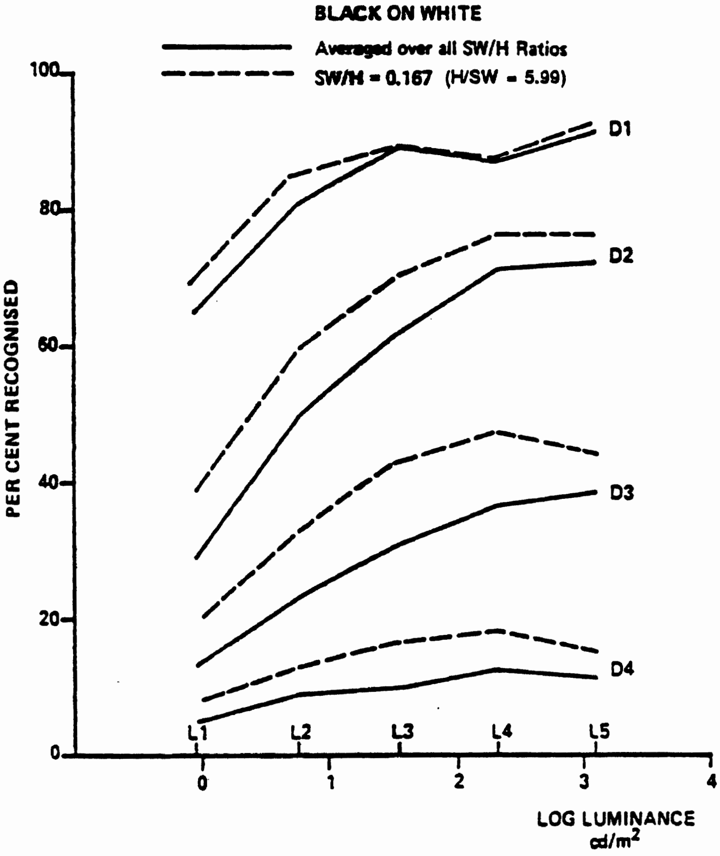


Figure 51. Effect of Luminance and Distance on the Recognition of Black Numerals on White Background.

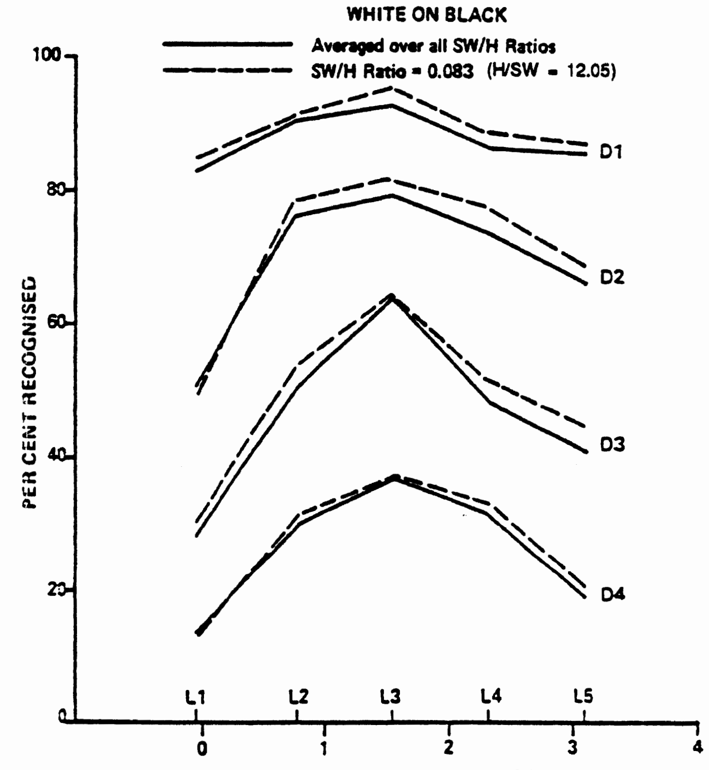


Figure 52. Effects of Luminance and Distance on the Recognition of White Numerals on Black Backgrounds.

luminance led to increases in percent recognition. Further, the optimal legibility for black characters was obtained for height to strokewidth ratios greater than 6:1. For white characters (Figure 52), increases in luminance increased percent recognition up to an optimal luminance of 34.3 cd/m^2 (10 fL). Further, the optimal legibility for white characters was obtained for height to strokewidth ratios greater than 12:1.

Other overall results include the following. White numerals on black backgrounds were more legible than black numerals on white backgrounds, except at high luminance. However, when averaged over all luminances, distances, numerals, and strokewidth to height ratios, white letters on black backgrounds were more often recognized than black numerals on white backgrounds. In terms of ability to recognize, straight line numerals performed much better than curved numerals. Further, legibility varies more with changes in visual angle than with changes in luminance.

Based on the results of this experiment, superior sets of black-on-white and white-on-black numerals were developed, as shown in Figures 53 and 54.

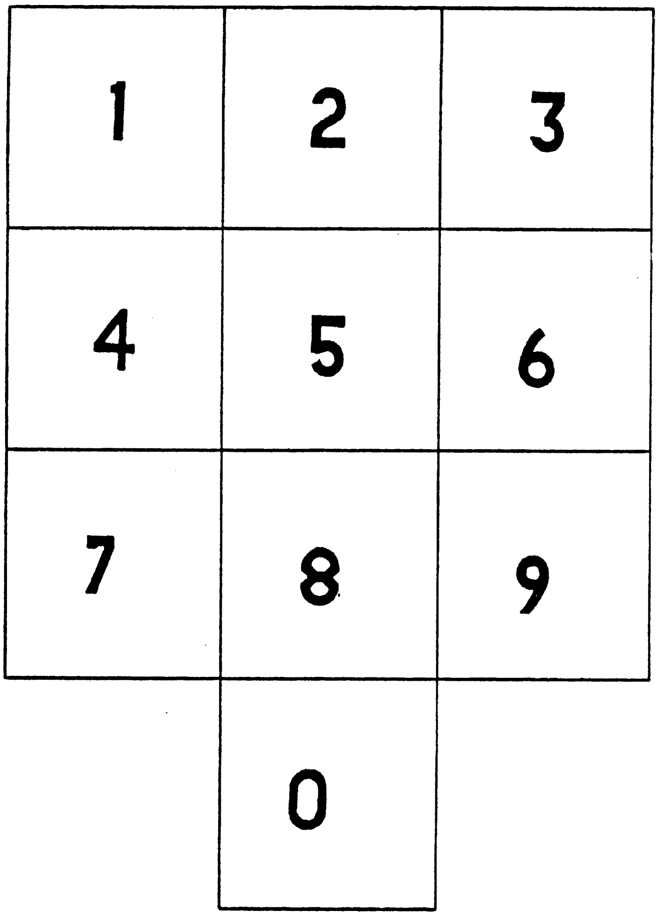


Figure 53. Improved Numeral Set for Black-on-white Numerals.

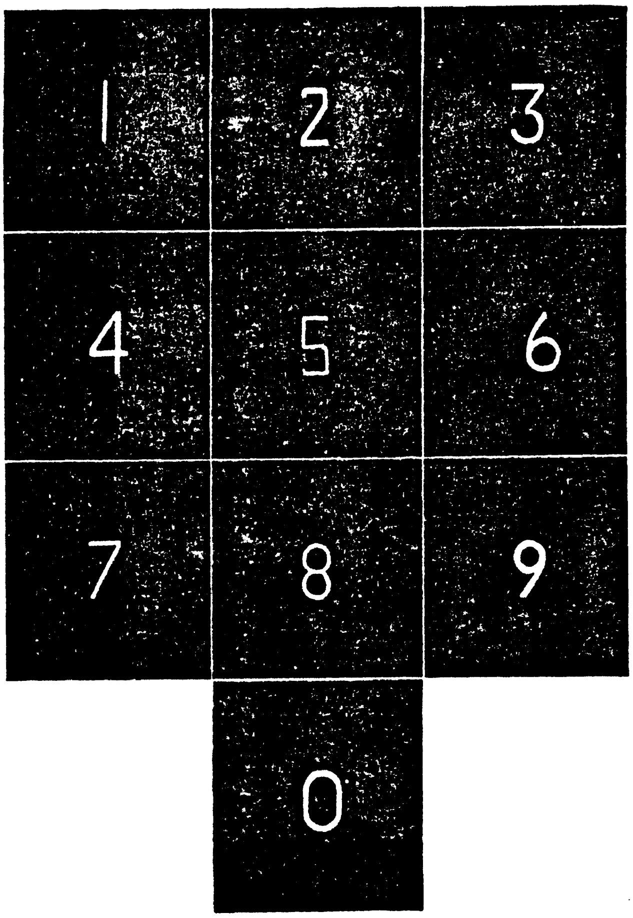


Figure 54. Improved Numeral Set for White-on-black Numerals.

Forbes, Saari, Greenwood, Goldblatt, and Hill (1976)

Two sets of experiments were performed. In the first, colored slides were projected in a darkroom. Two series of blank signs in seven unspecified colors and two series of target signs carrying a capital letter C or O with different orientations were presented. The height to width ratio of the letters was 1.25:1 and the height to strokewidth ratio was 5:1. Neutral overlays were used to vary the slide luminance between five levels simulating rural, suburban, and lighted city conditions (1.3 to 15 cd/m^2, 0.37 to 4.45 fL). An unreported number of people were asked to identify the letter or color shown. The exact nature of their response is unclear, as are the specifics concerning the test conditions.

Legibility distance (ft/in of character height) was found to be proportional to the logarithm of the luminance level of the brighter of each letter-background pair. The results for five color combinations are shown in Figure 55.

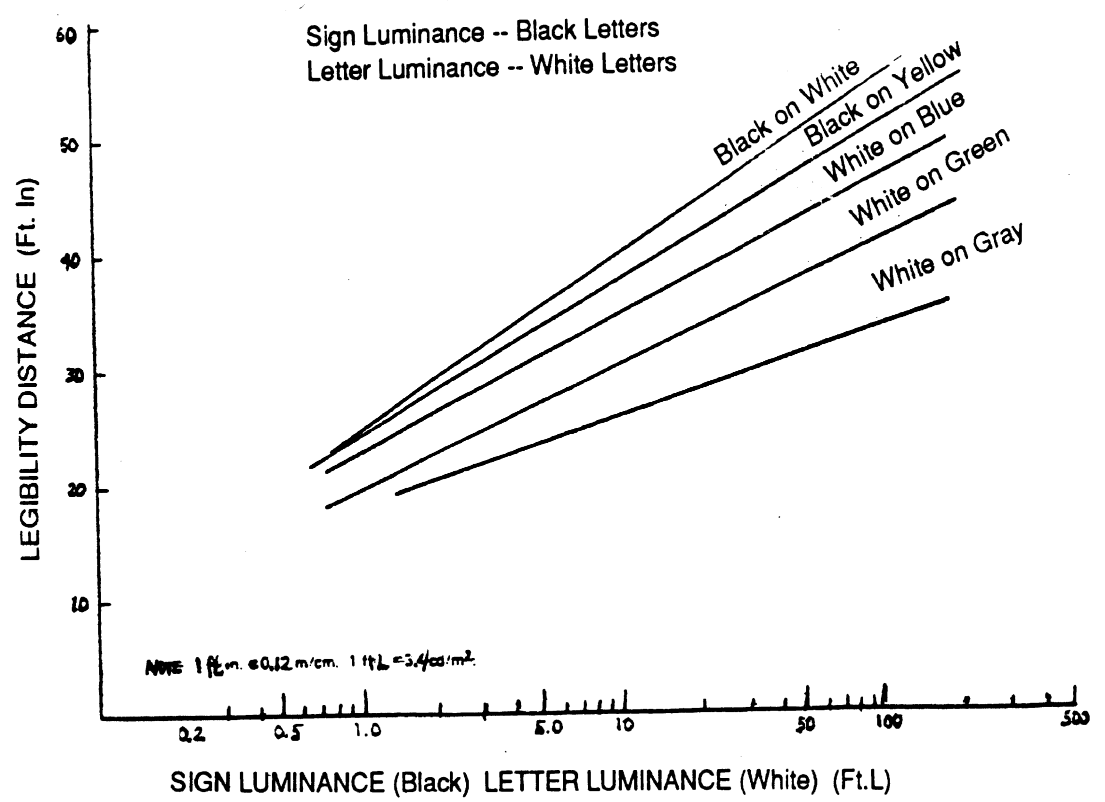


Figure 55. Average Legibility and Luminance for All Five Studies and Color Combinations.

The second series of experiments was conducted outdoors at night. Signs with 12-inch letters were viewed by 50 people. Letters were square E's with a height to strokewidth ratio of 5:1 or 7:1. Two color combinations (white on green, black on yellow) and two reflective materials were examined. Groups of six participants approached the signs from a distance of 1200 feet in a vehicle whose headlights were either in the low or high-beam position. Participants said when they could read the letter. Three ambient luminance levels (9.4, 1.2, and 0.034 cd/m^2 (2.75, 0.34, and 0.01 fL)) were used.

As in the laboratory experiment, the authors found a linear relationship between the logarithm of the sign luminance and legibility distance (as can be seen Figure 56), though the outdoor data seem to be noisier. Also having a major effect was the luminance ratio, with legibility distance increasing about 20% as luminance ratio increased from 2:1 to 10:1. The ambient illumination level had no effect on legibility distance.

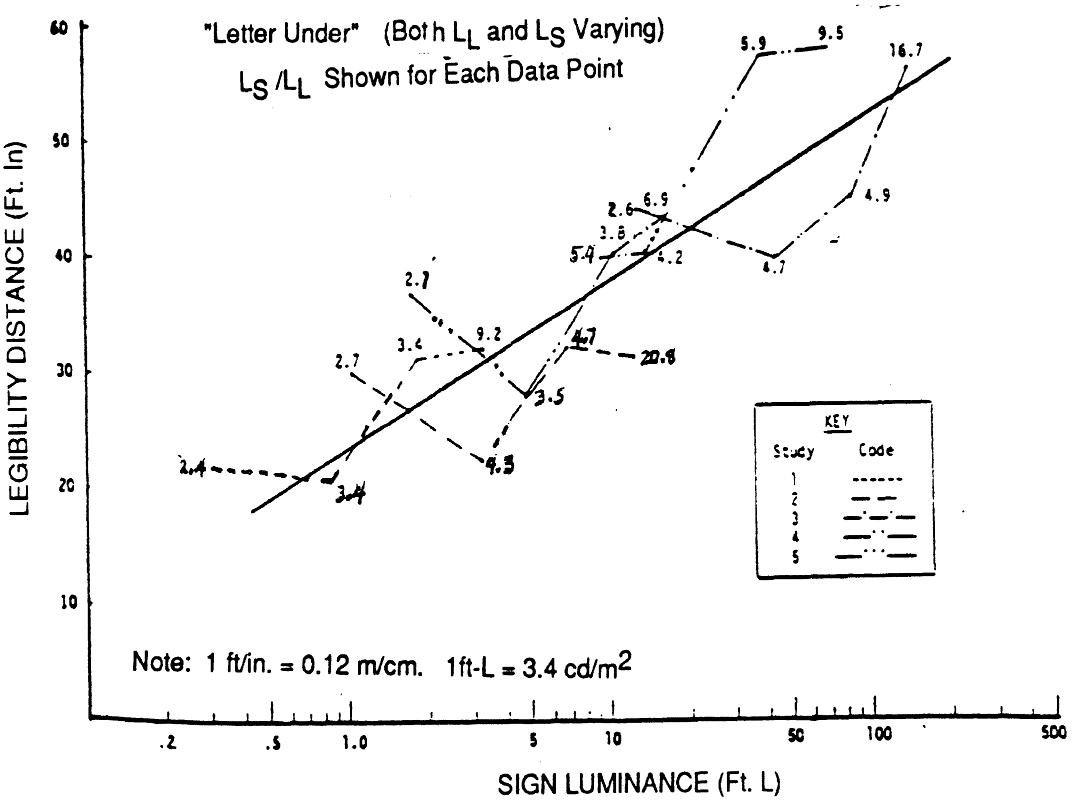


Figure 56. Legibility and Sign Luminance for Black Letters on a Yellow Background for All Five Studies and Contrasts.

Welsh, Rasmussen, and Vaughan (1977)

The authors collected readability data from 12 people in each of three age groups (20-25, 40-45, and 60-65 years old). Participants viewed black alphanumeric characters (eight-point Futura Demi-bold) against five levels of grey background. The 5 figure-to-background contrast ratios examined were 3.2:1, 6.6:1, 9.3:1, 12.8:1, and 16.2:1. Minimum illuminance required to identify all contrast combinations was determined at a viewing distance of 40 cm (15.7 in) under dim white and red illumination. Participants identified all characters while viewing them through an artificial pupil (2.0 mm).

As can be seen in Figures 57 and 58, the data indicate that a significant increase in illumination was required for successive decreases in contrast ratio for all age groups and under both illumination modes. Under red illumination, threshold luminance values showed a significant trend with age for all five contrast levels. Under white illumination, significant trends were indicated for three of the five contrast levels.

With reference to the younger group, individuals in the middle-aged and older groups required an average luminance increase of 18 and 63 percent respectively for equivalent readability scores under white illumination. Under red lighting, corresponding values were 18 and 58 percent.

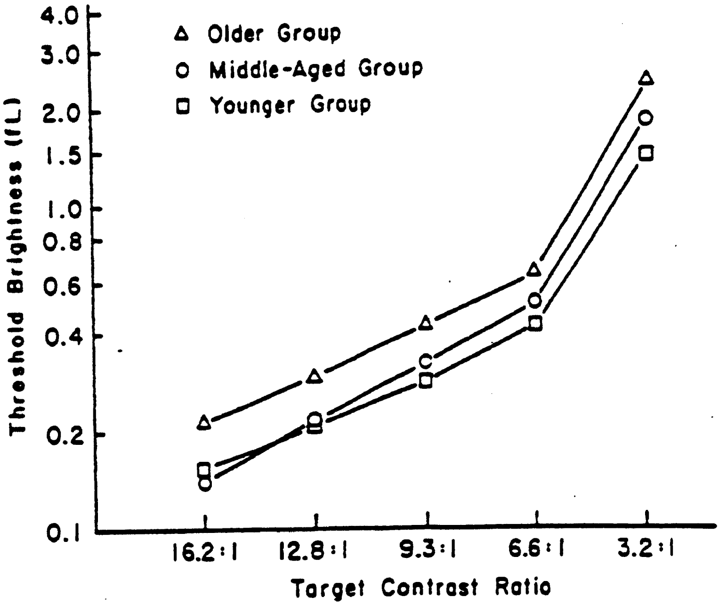


Figure 57. Threshold Luminance Under White Illumination as a Function of Target Contrast Ratio and Age.

Olson and Bernstein (1979)

A laboratory study measuring reaction time was carried out to define the effects of luminance, contrast, color, and driver visual characteristics on sign legibility distance. The study consisted of people identifying the orientation of five different sized Landolt rings at five distances (3.6, 4.8, 6.0, 7.2, and 8.4 m/cm letter height) for signs consisting of seven background colors (green, blue, red, black, yellow, orange, and white) and background luminances ranging from 0.07 to 211 cd/m^2 (0.02 to 62 fL).

For white legend signs, background luminance was fixed and the legend luminance was varied to find the zone from zero information transmission to the point where error-free performance was achieved for both high and low initial legend luminance. For black legends, background luminance was varied using the above procedure. Data were collected under low and high surround luminance conditions. Concurrently, a computer model was developed which could predict the legibility distance of a sign based on the laboratory data as well as geometric and photometric variables.

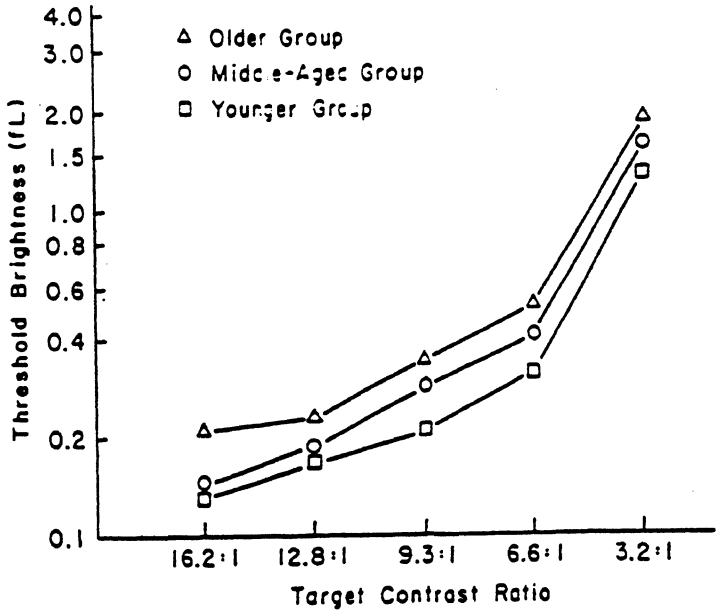


Figure 58. Threshold Luminance Under Red Illumination as a Function of Target Contrast Ratio and Age.

A two-step field study was then conducted at night using 18 people in which legibility distance predicted by the model was compared with legibility distance measured on a number of real and simulated signs. In one step, participants identified the orientation of the letter E on custom-made signs consisting of three different background reflectivities, four different legend reflectivities, and three different letter sizes. The second task had participants identify various freeway signs while traveling down the freeway at approximately 90 km/h. Since real highway signs were used, this study did not use a variety of colored backgrounds. In general, predicted legibility distances were within 10% of measured legibility distances found during the field study.

Results concerning effect of background color on legibility are as follows. At background luminance levels greater than 3.77 cd/m^2 (1.1 fL), red and green legibility data compare well. In terms of legibility, blue and green backgrounds also performed almost equally as well. In addition, white, yellow, and orange backgrounds produced similar legibility results, and had maximum legibility for luminances in the range of 3.4 to 34 cd/m^2 (1 to 9.9 fL).

It appears that color luminance contrast effects become inoperative at levels between 0.3 and 0.33 cd/m^2 (0.09 and 0.096 fL) for the ambient conditions used in this experiment. Further, data for black backgrounds were generally similar to those for colored backgrounds at the lowest luminance level tested.

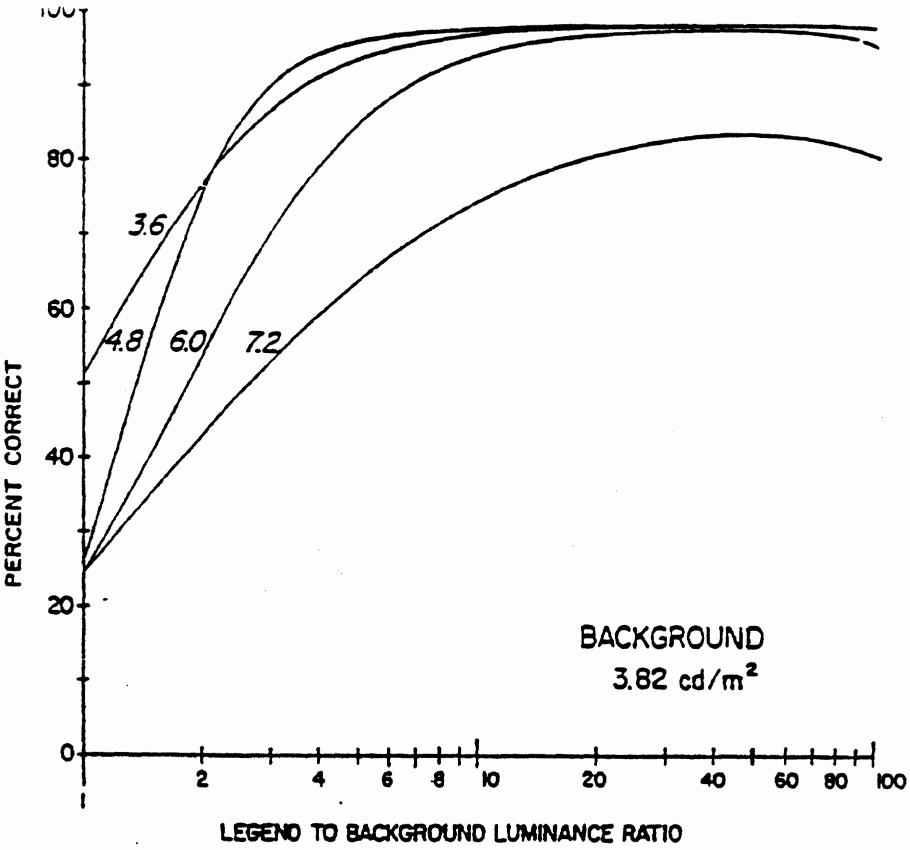
Olson, Sivak, and Egan (1983)

This article describes several studies concerned with the nighttime legibility of retroreflective signs. The first study was a laboratory investigation of the effect of sign background luminance, legend luminance contrast, surround luminance, background color, glare illuminance and angle, and a person's age on sign legibility. An unreported number of young and old people indicated the direction of a Landolt ring gap; different size rings simulated different viewing distances. Sign background color variations included green, blue, red, and black, all of which used white legends. White, yellow, and orange backgrounds were tested with a black legend. Sign legend luminance of the white legend varied from 0.038 to 733.0 cd/m^2 (0.01 to 214 fL). Surround luminance varied between three levels: 0.03, 3.43, and 17.0 cd/m^2 (0.009, 1, and 5 fL).

As can be seen in Figure 59, the results indicate that legend luminance contrast is the most important variable in sign legibility at night, and that there is a relatively narrow range of optimum contrast. Further, maximum legibility is achieved at a contrast of 30 to 60:1.

As a rough rule of thumb, the data suggest that changing background luminance by a factor of ten requires that the contrast ratio be changed by a factor of two to maintain a given legibility level. Sign background color and surround luminance both have relatively minor effects on legibility. In general, the older participants did much poorer than the younger participants. However, the difference was minimized by high-luminance backgrounds.

The next two studies were carried out under field conditions. Small signs were used, and 12 people rode toward them, pressing a button to indicate the distance at which they became legible. The results indicate that the expected age effect on legibility of signs is eliminated if the age groups are matched in terms of their low-luminance/high-contrast visual acuity.



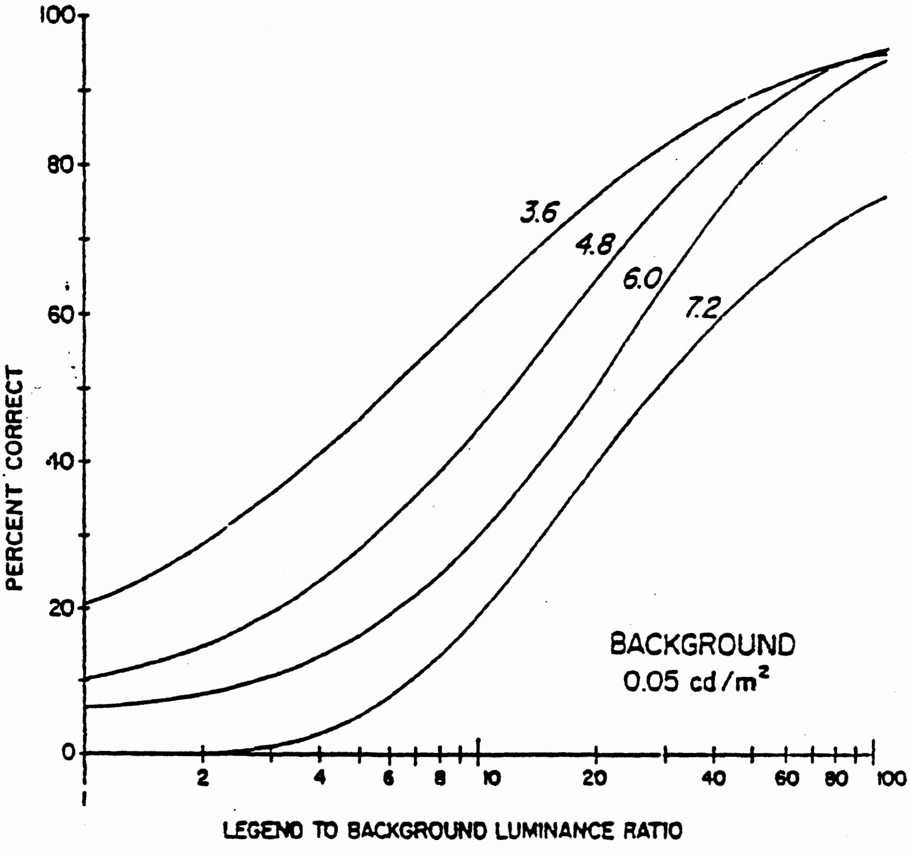


Figure 59. Percent Correct Responses to Four Legend Sizes as a Function of Legend Luminance Contrast and Background Luminance.

Sivak and Olson (1985)

The authors of this study present optimal and minimal sign luminance recommendations based on a review of available applied research. To account for various conditions, the geometric mean was computed from various studies' optimal luminance values, as can be seen in Table 14.

Table 14. Minimal Sign Luminance Recommendations from Various Studies.

Luminance (cd/m^2)	Study Characteristics
343.0	Allen and Straub laboratory study - an asymptote was apparently not reached even with the highest level tested.
34.3	Allen et al. field study - dark rural (used both 100% and 75% legend/background luminance contrast).
60.0	Dahlstedt field study.
206.0	Hind et al. laboratory study - the data appear to asymptote at 206 cd/m^2 (60 fL).
55.0	Olson et al. laboratory study - (recommended luminance: 10 - 100 cd/m^2, 2.9 - 29 fL).
24.0	Smyth laboratory study.

Optimal recommendations are based largely on peak luminance-legibility relationships. Specifically, 75 cd/m^2 (22 fL) is recommended as the optimal luminance for signs with light (white, orange, or yellow) backgrounds and black legends. The authors suggest 12:1 as the optimal legend to background contrast. In the absence of other criteria, minimal recommendations are based on performance levels of 6 m/cm letter height (0.72 ft/in) (20/23 visual acuity) for younger persons and 4.8 m/cm (0.576 ft/in) (20/29 visual acuity) for older persons. The recommended minimum luminance of the lighter components is 2.4 cd/m^2 (0.7 fL). This recommendation applies to light backgrounds (white, yellow, or orange) with black legends, and to white legends with dark (green, blue, red, or brown) backgrounds having a background luminance of up to 0.4 cd/m^2 (0.12 fL).

Summary

Legibility Distance: Uhlaner (1941) proposed a formula in which legibility distance is a parabolic function of a character's strokewidth to height ratio. Forbes et al. (1976) found in a study of questionable methods that legibility distance was proportional to the logarithm of the luminance level of the brighter of a letter/background pair. However, no letter height prediction equation was proposed.

Letter Brightness: Smyth (1947) found an approximately linear relationship between ideal letter brightness and the logarithm of the background brightness. Further, he found a range of mean ideal letter brightnesses of 75 lux (7 fc) for a 0.015 lux (0.0014 fc) background brightness to 269 lux (25 fc) for background brightness of 15 lux (1.4 fc). Hind, Tritt, and Hoffman (1976) found that for black characters, increases in luminance led to increases in percent recognition, while for white characters, increases in luminance increased percent recognition up to an optimal luminance of 34.3 cd/m^2 (10 fL). Based on an accumulation of previous data, Sivak and Olson (1985) recommend 75 cd/m^2 (22 fL) as the optimal luminance for retroreflective signs with light backgrounds and black legends.

Other Key Conclusions from this Topic

- Optimal height to strokewidth ratio for black characters was between two and three times greater than for white characters.

- After averaging over many legibility factors, it was found that legibility was best for white-on-black characters, except under conditions of high luminance. Differences between white-on-black and vice-versa tend to be quite small.

- Increases in legibility obtained by changing from a reasonably good font to a hypothesized better font are small relative to changes in other factors affecting legibility, such as contrast, luminance, and illuminance.

- Readability data differed significantly only between old (60-65 years old) and young people (20-25 years old). The readability data for middle-aged individuals (40-45 years old) were similar to that of young people.

- A rough rule of thumb is that changing background luminance by a factor of ten requires a change in contrast ratio by a factor of two in order to maintain a given level of legibility.

- In general, visibility performance measures are related to the logarithm of lighting variables, such as illumination, contrast, brightness, etc.

AIRCRAFT AND MILITARY APPLICATIONS

This section describes several studies whose purpose was to develop or identify design recommendations for character heights, height to width ratios, illumination and luminance levels for aircraft displays. These recommendations are not aircraft-specific; they have general applicability to all displays.

Brown (1953)

Two experiments described by Brown (1953) were designed to determine the effects of height and height-to-width (H:W) ratio upon the legibility of capital letters used in aircraft cockpit displays. The ensuing description is common to both sets of experiments. Height to strokewidth remained constant at 6:1. Five levels of red transillumination simulating nighttime and two levels of illumination simulating daylight were tested. The transillumination conditions were presented for two exposure durations (40 msec and 200 msec), whereas the daylight conditions were exposed for 7 msec (which seems unreasonably brief). For each duration condition, after receiving "several" example exposures, 20 subjects read aloud slides with groups of three letters (25 groups per condition). All participants were Naval enlisted men, totaling 120 for the entire study.

In the first set of experiments, letter height was held constant (.156 inches) while H:W ratios varied from 1:1 to 1:.55 in four increments. In the second set of experiments, letter H:W ratio was held constant at 1:1 while six heights were tested, ranging from .12 to .18 inches.

Results are presented in tables containing number of and percent errors, as well as graphs representing the same information. No statistical tests were performed on the data (e.g., ANOVA) since these tests were not yet in widespread use at the time of the report.

Figures 60 through 62 apply to the first set of experiments. In general, according to Figure 60, a H:W ratio of 1:1 seems optimum. This holds true for all luminance levels. Luminance had only a minor effect on performance for the range examined. In Figure 61, luminance had a major effect for short exposure durations. There does not seem to be any interaction with H:W ratio. Under daylight conditions (Figure 62), an illumination level of 80 fc proved superior over 40 fc for every height to width ratio. Again, a H:W ratio of 1:1 seems best.

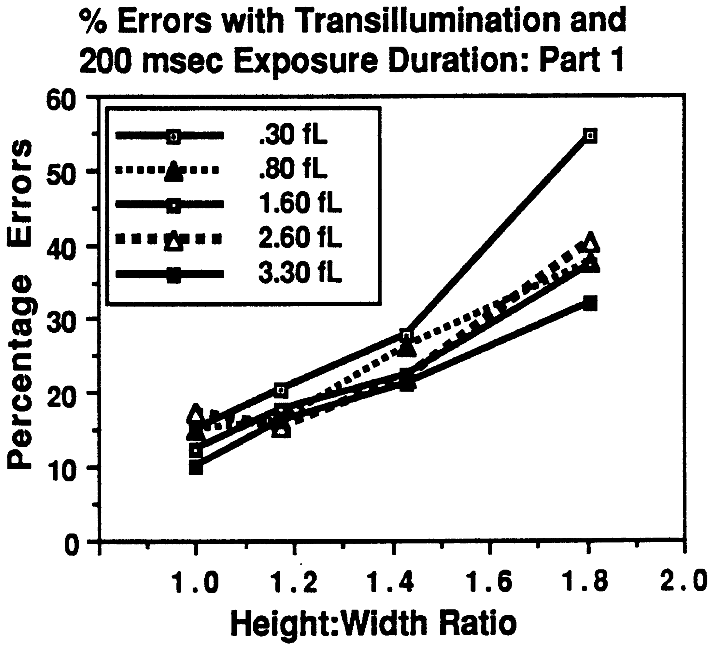


Figure 60. Percent Errors with Transillumination and 200 msec Exposure Duration--Part 1.

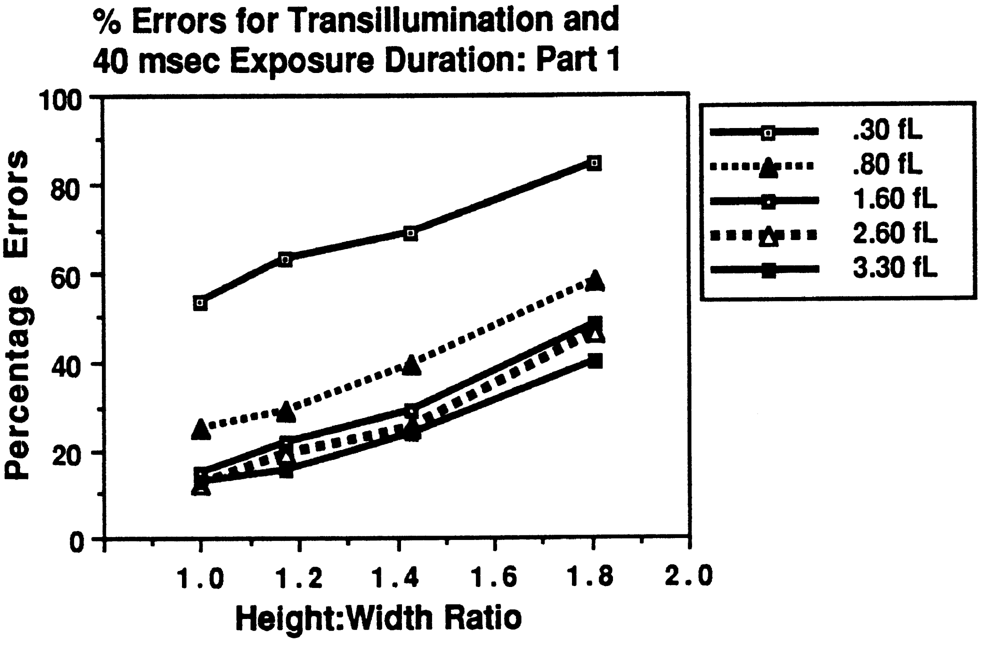


Figure 61. Percent Errors with Transillumination and 40 msec Exposure Duration--Part 1.

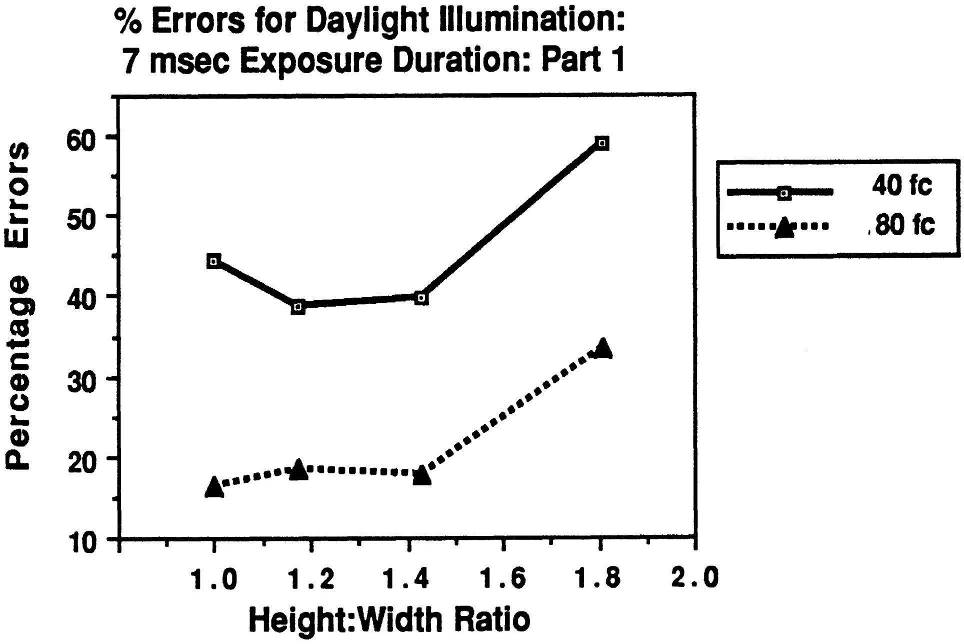


Figure 62. Percent Errors with Daylight Illumination and 7 msec Exposure Duration--Part 1.

Figures 63 through 65 apply to the second set of experiments. As Figure 60 illustrates, there was a linear relationship between height and percent errors. Where an increase in luminance by an order of magnitude reduces errors by 5%, increasing height from .12" to .18" reduces them by 10%. If this trend continues, one could expect error-free performance at a height of .25". Figures 63 through 65 show similar trends.

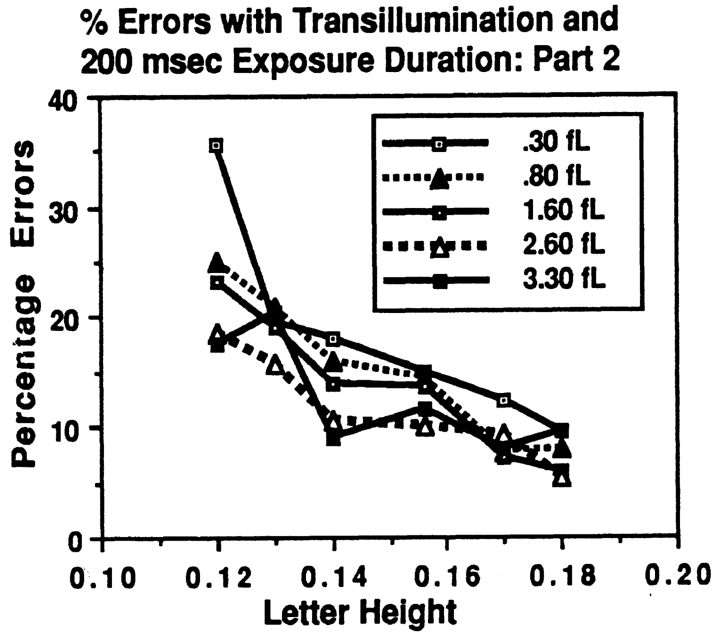


Figure 63. Percent Errors with Transillumination and 200 msec Dxposure Duration--Part 2.

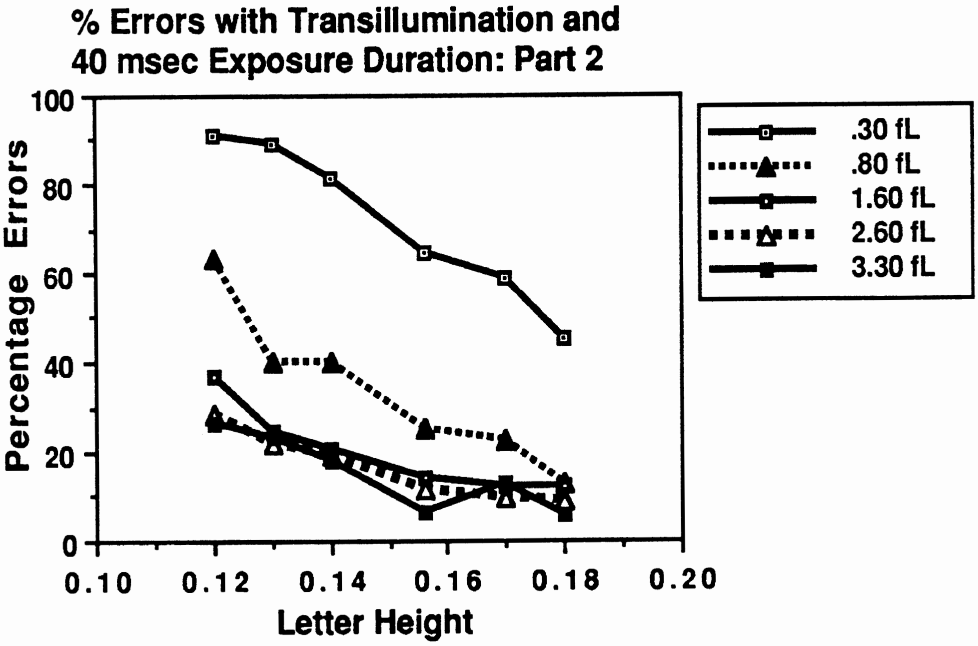


Figure 64. Percent Errors with Transillumination and 40 msec Exposure Duration--Part 2.

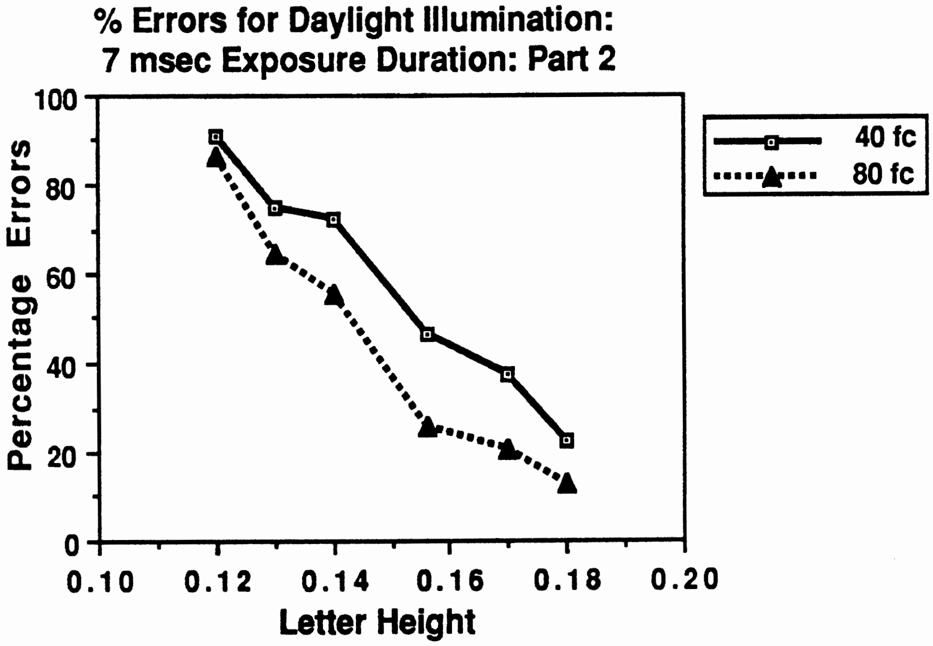


Figure 65. Percent Errors with Daylight Illumination and 7 msec Exposure Duration--Part 2.

In general, the authors conclude that for uniform strokewidth capital letters, legibility improves with increasing width up to a 1:1 height to width ratio when height was constant at .156 inches. Marked loss in legibility is found when letter width is narrower than 2/3 the height. An improvement in the experimental design would be to test height to width ratios past the 1:1 ratio (i.e., wider letters) in order to definitively find the optimum ratio. Since errors were still decreasing at the 1:1 ratio, a "floor" had not been determined.

A problem with the methods employed in this study is that within each condition, each participant viewed the slides in exactly the same order. This lack of randomization could possibly lead to confounded results. Another noteworthy point is that the author claims a 1.43:1 height to width ratio is in the "acceptable range." However, subjects made between 20% and 30% errors with that condition, which often cannot be considered acceptable.

U.S. Department of Defense (1981a,b) -- Mil-Std-1472C and Mil-Hdbk-759A

These two Department of Defense human factors design documents provide useful guidelines concerning legibility of visual displays. According to Van Cott and Kinkade (1972), in particular pages 88-89, it appears that these guidelines were based upon the work of Brown and others, some of which was described previously. Although the recommendations are derived from research done on aircraft, they have widespread applicability for all types of displays. Following are examples taken from section 5.2 (Visual Displays) and 5.5 (Labeling) of Mil-Std-1472C.

With respect to illumination requirements, when dark adaptation is minimal, use low brightness, adjustable white light. Where complete dark adaptation is required, use low luminance (.07 - .34 cd/m2) red light (greater than 620 nm). Automotive lighting levels are rarely so low that full dark adaptation occurs. Furthermore, there is controversy over the merits of the red light recommendation, a recommendation that may be changed in future versions of the Standard.

Luminance of transilluminated displays should be at least 10% greater than the surrounding luminance. Where glare must be reduced, this luminance should not exceed 300%. When displays will be used under varied ambient illumination, a variable display brightness control should be provided, in order to cover the full range necessary for legibility in all conditions.

With respect to design of characters on a label, 1472 specifies letter and numeral widths to be 3/5 of the height, with the exception of "W", "M" and "4" which should be 4/5, and "I" and "1" which should be 1/5. The standard also provides a useful table of label size versus luminance, reproduced on the following page (Table 15).

Another useful table in the Mil-Standard presents character height requirements for given viewing distances. (See Table 16.)

Table 15. Label Size Versus Luminance.

Markings	Height* 3.5 cd/m^2 (1 fL) or BELOW	ABOVE 3.5 cd/m^2 (1 fL)
For critical markings, with position **variable** (e.g., numbers on counters and settable or moving scales)	5-8 mm (0.20-0.31 in)	3-5 mm (0.12-0.20 in)
For critical markings, with position **fixed** (e.g., numerals on fixed scales, controls, switch markings, or emergency instructions)	4-8 mm (0.16-0.31 in)	2.5-5 mm (0.10-0.20 in)
For **non-critical** markings (e.g., identification labels, routine instructions, or those required only for familiarization)	1.3-5 mm (0.05-0.20 in)	1.3-5 mm (0.05-0.20 in)

* Values assume a 710 mm (28 in) viewing distance, typical for viewing a display. For a distance, D, other than this, multiply the above values by D/710 mm (D/28 in).

Source: U.S. Department of Defense (1981b)

Table 16. Character Height for Various Viewing Distances.

Viewing Distance	Minimum Height
Less than 500 mm (19.7 in)	2.3 mm (0.09 in)
0.5-1.0 m (19.7-39.4 in)	4.7 mm (0.18 in)
1.0-2.0 m (39.4-78.7 in)	9.4 mm (0.37 in)
2.0-4.0 m (78.7-157.5 in)	19 mm (0.75 in)
4.0-8.0 m (157.5-315.5 in)	38 mm (1.50 in)

Source: U.S. Department of Defense (1981b)

Rogers, Spiker and Cicinelli (1986)

A series of three experiments investigated the effects of display luminance contrast on the legibility of self-luminous displays in aircraft cockpits at low, moderate, and high illumination levels. The studies avoided the usual confounding of the variables, and were conducted using luminance levels representative of operational conditions.

The first experiment (low ambient illumination) involved 33 people. After being given 10 minutes to dark adapt, participants verbally identified breaks in green Landolt rings. Rings were viewed at 4.27 meters (14 ft), not a typical viewing distance for instrument panels. The background was varied to give contrasts of 2:1, 4:1, and 8:1. Luminance levels of 13.7, 41.1, and 82.2 cd/m2 were examined. Participants adjusted symbol size to meet two conditions: threshold level legibility and comfort level legibility.

An ANOVA showed the effect of luminance contrast was statistically significant in both threshold and comfort cases. Symbol luminance was not significant. The 2:1 contrast condition differed significantly from the 4:1 and 8:1, but the 4:1 and 8:1 did not differ significantly from each other (though the size required was slightly smaller). These findings suggest that under low ambient illumination, increasing the contrast ratio above 4:1 will have a minimal effect on detection performance, especially for self-paced tasks. Table 17 shows that threshold legibility ranged from 5.7 to 7.8 minutes of arc, while comfort legibility was 1.5 times threshold legibility.

Table 17. Mean Symbol Subtenses (Arcmin) for Legibility in Experiment 1.

	Contrast Ratio		
	2:1	4:1	8:1
Threshold legibility	7.8	6.0	5.7
Comfort legibility	11.5	9.5	9.2

Source: Rogers, Spiker and Cicinelli (1986)

The second experiment dealt with luminance contrast under moderate ambient illumination. The experiment compared random-scan and raster-scan CRT displays under overhead lighting of 108 lx (10 fc). Table 18 shows the various viewing conditions examined, "Illumination" referring to the illumination on the display screen.

For each condition, 10 people were pushed in a moveable chair toward green Landolt rings from a distance of 7.9 meters (26 ft) until first threshold and then comfort legibility were achieved. Results show the random-scan CRT to be superior for both threshold and comfort cases in both the 1.2:1 and 1.5:1 contrast ratio conditions, which is to be expected. Symbol and background luminances had no significant effects on legibility apart from their contribution to the contrast ratio. This finding is consistent with that from the first experiment. Table 19 shows that the size ranges are similar to that found in the first experiment, and once again, comfort legibility was 1.5 times that of threshold.

Table 18. Viewing Conditions--Random-scan vs. Raster-scan CRT.

Condition	Display Mode	Illumination		Symbol	Bkgd lum	
		lx	(fc)	cd/m^2	(fL)	CR
1	Raster	29 062	(2700)	144:96	(42:280)	1·5:1
2	Raster	53 820	(5000)	267:192	(78:56)	1·4:1
3	Raster	107 639	(10 000)	497:418	(145:122)	1·2:1
4	Random	29 062	(2700)	144:96	(42:28)	1·5:1
5	Random	107 639	(10 000)	497:418	(145:122)	1·2:1
6	Random	29 062	(2700)	720:96	(210:28)	7·5:1
7	Random	29 062	(2700)	195:65	(57:19)	3:1
8	Random	53 820	(5000)	524:175	(153:51)	3:1
9	Random	107 639	(10 000)	1131:377	(330:110)	3:1

Source: Rogers, Spiker and Cicinelli (1986)

Table 19. Minimum and Comfort Thresholds (Arcmin) for the Two Scan Modes As a Function of Contrast Ratio.

	Raster Scan (Contrast)			Random Scan (Contrast)			
	1.2:1	1.4:1	1.5:1	1.2:1	1.5:1	3:1	7.5:1
Minimum	9.9	7.1	6.3	4.5	3.6	3.6	3.5
Comfort	13.4	9.0	7.0	5.9	4.5	4.6	4.3

Source: Rogers, Spiker and Cicinelli (1986)

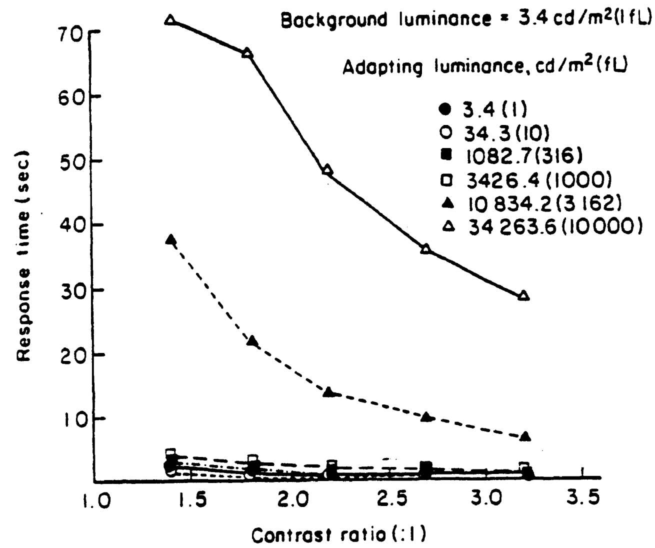
The third experiment (the most critical in terms of instrument panels) examined luminance contrast under high ambient illumination. This study incorporated the realism of having to adapt quickly to high levels of illumination, as is often required in aircraft cockpits. Ten subjects gazed into an adaptation field, and upon cue turned to the CRT screen and verbally identified the break in a Landolt ring. Table 20 gives the viewing conditions.

With respect to contrast ratio and adaptation luminance, all main effects were significant (background luminance, adaptation luminance, and contrast ratio) as well as their interactions. Results are presented in Figures 66 and 67. The graphs suggest that contrast ratio exerts little influence at low adaptation luminances, but a significant one at the higher luminances. An important point arises when applying these results to automotive displays rather than aircraft. Of the six adaptation levels tested, the two highest are not typically representative of the road scenes a driver may encounter, whereas the other four conditions are. This could imply that it is not critical to address adaptation in an automotive display study, for the results here show no significant difference between the levels. However, during bright, snowy weather the outdoor environment and the instrument panel could differ by two orders of magnitude. Thus, although this condition is uncommon, adaptation is a potential issue for future instrument panel studies.

Table 20. Viewing Conditions--Luminance Contrast Under High Ambient Illumination.

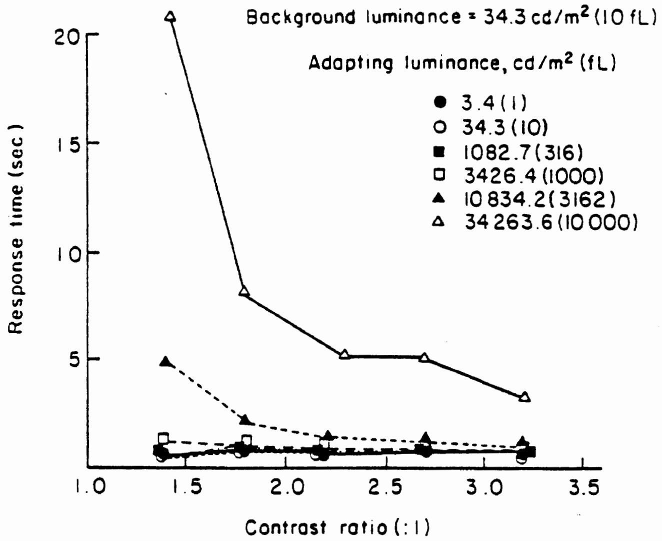
Type of design	Contrast ratio	Display background luminance cd/m²	(fL)	Adaptation luminance cd/m²	(fL)	Typical equivalent
				3·4	(1)	Snow in full moon
	1·4:1			34·3	(10)	TV screen
	1·8:1	3·4	(1)	1083	(316)	Average sky on cloudy day
Factorial	2·2:1	34·3	(10)	3426	(1000)	Average earth on clear day
	2·7:1			10 834	(3162)	Average sky on clear day
	3·2:1			34 264	(10 000)	Illumination from sun
Non-factorial	4:1, 6:1 10:1, 15:1 20:1, 30:1	3·4	(1)	34 264	(10 000)	Illumination from sun
Non-factorial	1·4:1 2:2:1 3:2:1	257	(75)	34 264	(10 000)	Illumination from sun

Source: Rogers, Spiker and Cicinelli (1986)



Source:
Rogers, Spiker, and Cincinelli (1986)

Figure 66. Effects of Adaptation Luminance and Contrast Ratio on Mean Response Time for a 3.4 cd/m2 (1 fL) Background.



Source: Rogers, Spiker and Cicinelli (1986)

Figure 67. Effects of Adaptation Luminance and Contrast Ratio on Mean Response Time for a 34.3 cd/m2 (10 fL) Background.

The effects of adaptation "mismatch" (background luminance vs. adaptation luminance) is illustrated in Figure 68. Neither the 1.4:1 nor the 3.2:1 contrast ratio curve rises significantly until a log ratio of two is reached. Therefore, recovery is almost immediate as long as the background and adaptation luminances do not differ by more than a factor of 100. When they do differ by this much, response times of up to 720 msec (1.16 min) were seen, which is dangerously long to look away from the road while driving.

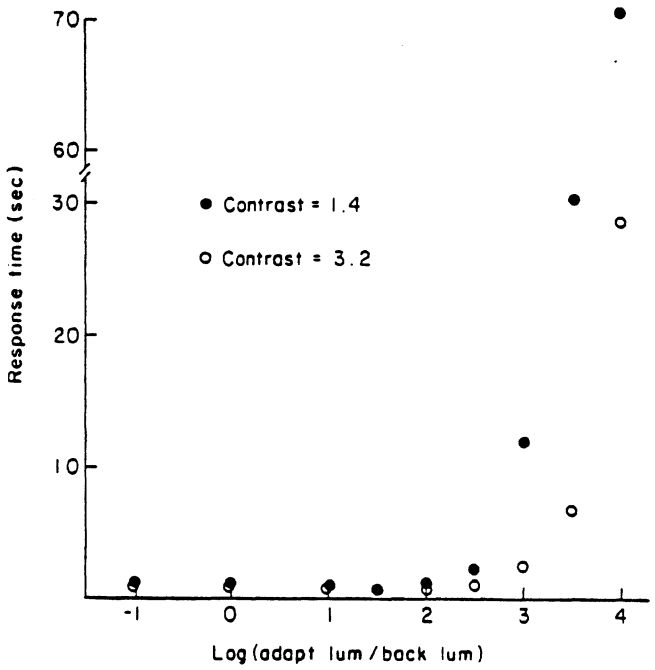
As part of this project, the authors developed several models for response time as a function of the display parameters. Three equations evolved:

$RT = 100(CR^{-1})$ for $3.4\ cd/m^2$ (1fL)

$RT = 31(CR^{-2})$ for $34.3\ cd/m^2$ (10 fL)

$RT = 8(CR^{-2})$ for $257\ cd/m^2$ (75 fL)

where: RT = Response Time
CR = Contrast Ratio



Source: Rogers, Spiker and Cicinelli (1986)

Figure 68. Effects of Adaptation "Mismatch" on Mean Response Time at Contrast Ratios of 1.4 to 1 and 3.2 to 1.

Further, the authors identify a more general predictive model for response time (below). However, they state that collecting the data and solving for the constants is a matter for further study.

$$RT = k(AL)^{a}(BL)^{-b}(CR)^{-c}$$

where:
RT = Response Time
AL = Adaptation Luminance
BL = Background Luminance
CR = Contrast Ratio
k, a, b, c = undetermined constants

Summary

The critical documents in this section are the Department of Defense standard and handbook. They explicitly specify the minimum sizes required for labels. While some may view these requirements as only applying to weapon systems, Mil Standard 1472C is widely accepted as the human factors Bible, and failure to meet its very conservative minimums is a serious oversight.

Also of interest in this section is the Rogers et al. study. Of particular note are the extremely long response times associated with conditions where the adaptation luminance is two orders of magnitude above the display illumination level. While those display conditions are not common in automobiles, they may occur when sunlight is reflected off of snow. Efforts should be made to obtain vehicle interior and exterior illumination levels for those conditions. Bos and Kerst (1988) examined this issue under a few conditions, but it is just a start. Much more work remains to be done.

LEGIBILITY PREDICTIONS FOR TEXT ON AUTOMOBILE INSTRUMENT PANELS

Presented in this section are two studies concerned specifically with the legibility of instrument panel displays. Older drivers, especially those over 45 years old, have more difficulty than younger drivers in reading instrument panel displays. To assure that all drivers can read the displays, design specifications should be slanted towards older drivers. The two studies included in this section concern problems involved when testing strictly older people and the development of recommendations for various legibility parameters to accommodate them.

Sauter and Kerchaert (1972)

Sauter and Kerchaert developed a procedure to estimate how well older drivers (i.e., worst case) could read instrument panel characters based on data from people of any age or visual acuity. In their experiment, 30 people over 45 years old and 10 under age 45 participated. A mockup vehicle with an instrument panel containing three rectangular openings was used. The side openings were 8.5 cm from the steering wheel center line. The center opening was 28 inches from the driver's eyes, and the steering wheel was removed. Stimuli displayed in these openings varied in size from 0.22 to 0.89 min^{-1}, and instrument panel illumination was either 21.5, 215, or 1075.3 lux (2, 20, or 100 fc).

The estimation procedure consisted of two parts. First, the response time to typical instrument panel characters displayed through the panel openings was found. Response time was found by having a person (fixating at a distant target) look down at one of the three instrument panel openings when prompted by a tone, recognize the character in the opening, and press a switch to stop a timer. Auditory reaction time (how long a person took to press a button after hearing a tone) was measured separately. Perception time was response time minus audible reaction time.

The second part of the estimation procedure involved measuring the visibility of alphanumeric characters used in the first part of the experiment using a Luckiesh-Moss visibility meter. This meter, used similarly to binoculars, has two adjustable neutral density filters in front of the observer's eyes to vary the visibility of a target stimuli. The filters reduce the illumination level while holding all other factors (character size, contrast ratio, etc.) constant. Positioned 28 inches from the center instrument panel opening, the filters in the meter were lightened until the "subject could first distinguish or read" the character displayed.

Using regression analysis, the authors reported the following two relationships:

Perception Time (Sec) = Tp = $(1.71/V^2) - (0.81/V) + 0.42$ (1)

where:

V = visibility meter reading in density units.

Response Time (Sec) = RT = $(1.58/V^2) - (0.55/V) + 0.56$ (2)

A plot of these equations produced by the authors of this review, including regression equations and correlation coefficients, can be found in Figures 69 and 70. It should be noted that the regression equations associated with these figures differ significantly from those generated by Sauter and Kerchaert. Exact reproduction of regression equations is difficult due to slight differences in computational procedures. However, the magnitude of the difference between the two regression equations raises questions as to the accuracy of the Sauter and Kerchaert equations. Because of this, additional analysis of the equation 2 data, the more relevant of the above two equations, was performed by the authors of this review and is presented below.

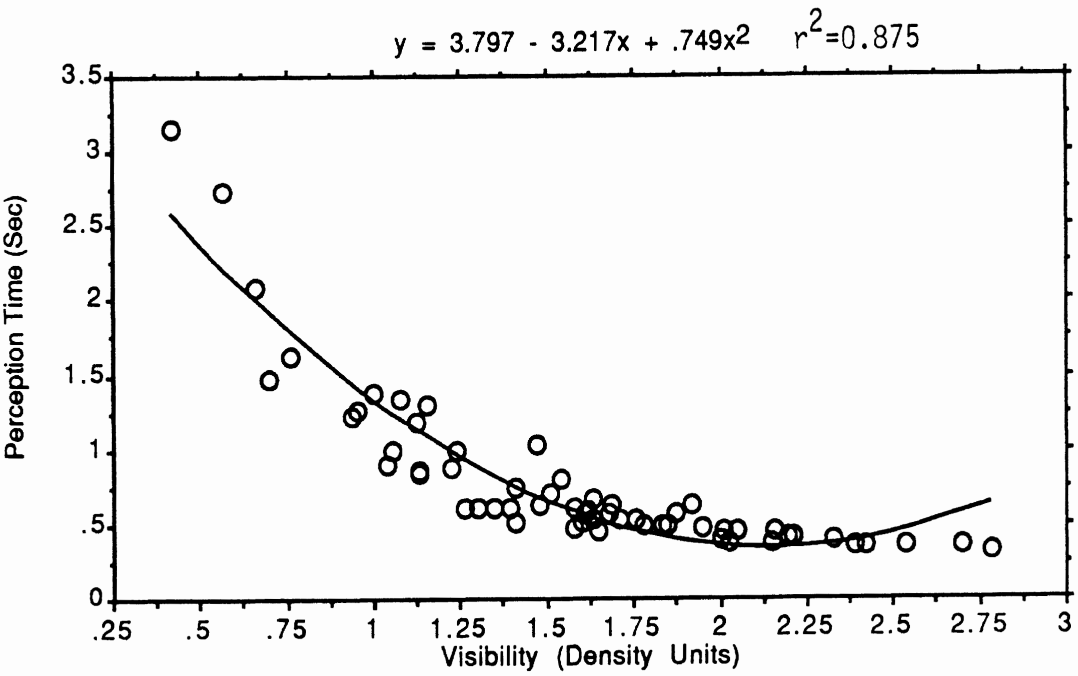


Figure 69. Perception Time for All Participants As a Function of Visibility (Quadratic Model).

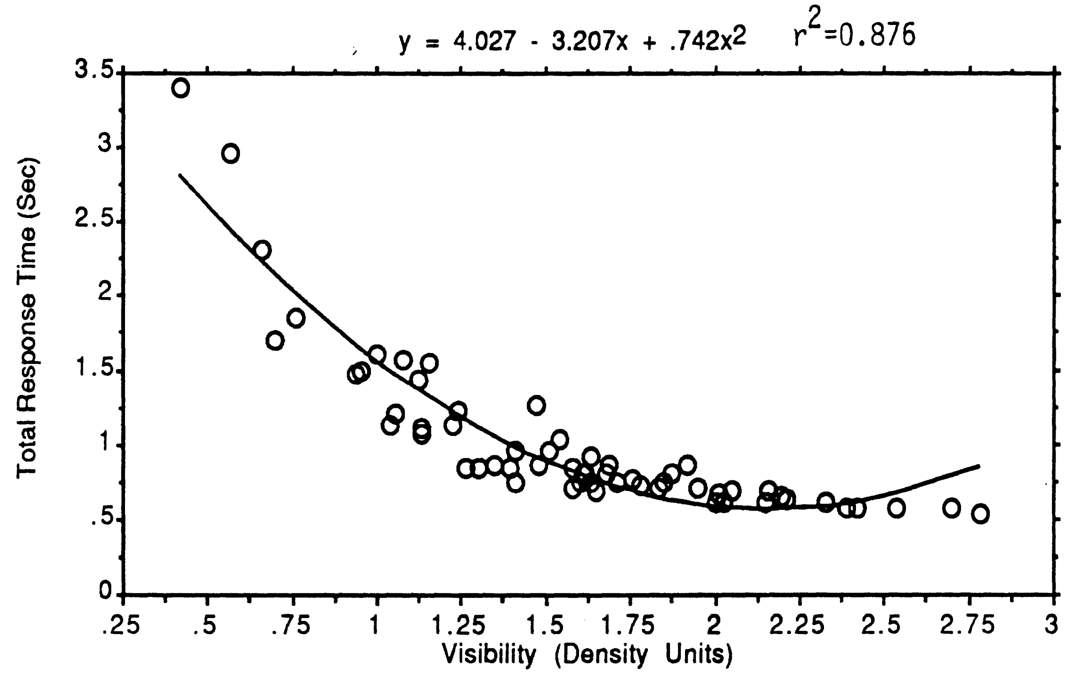


Figure 70. Response Time for All Participants As a Function of Visibility (Quadratic Model).

As can be seen in Figure 70, the parabolic regression line provides a good fit, except for the extremes, especially the lower extreme. The variability at this extreme can best be explained by a response time floor effect. At low visibility levels, the predominant task changes from reading the character to searching for a character, which also changes the relationship between reaction time and visibility. Therefore, for low visibility conditions, small changes in visibility density will produce large changes in reaction time different from those at higher visibility densities.

To provide a better fit to the extreme data points of Figure 70, a third order regression was developed:

$$RT\ (sec) = (-0.547/V^3)+(3.38/V^2)-(7.052/V)+5.665,\ r^2 = 0.929$$

A plot of this equation can be found in Figure 71. This model provides an increase of 0.053 in the correlation coefficient over the quadratic model, which is a large increase for correlations so close to 1. However, the authors of this report could find no theory or data to support a cubic relationship between visibility and reaction time, making the utility and accuracy of this equation for more generalized data questionable.

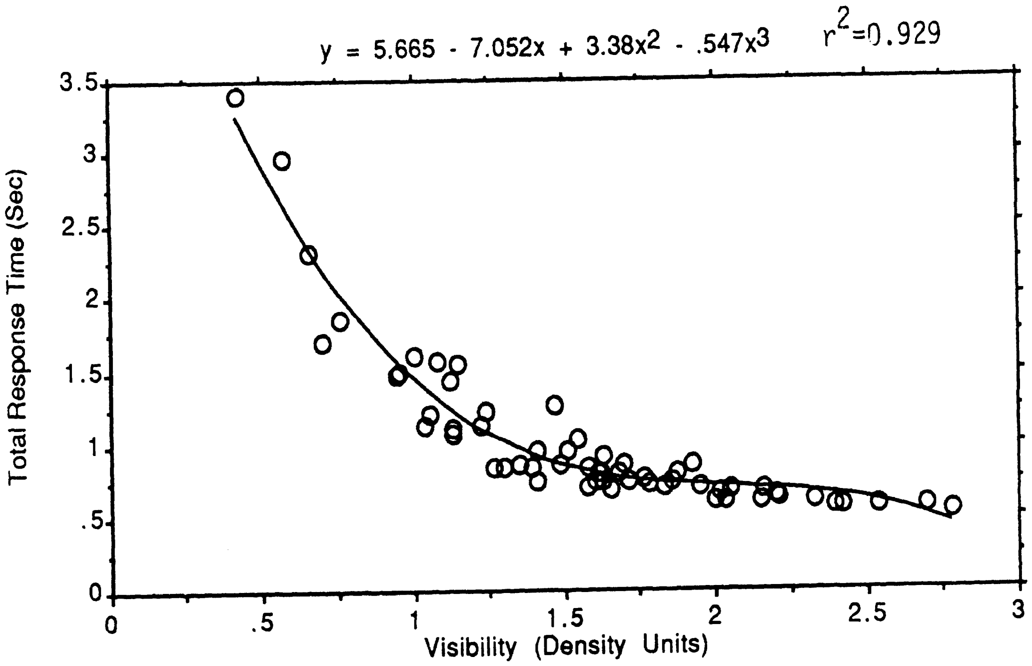


Figure 71. Response Time for All Participants As a Function of Visibility (Cubic Model).

As noted elsewhere in this report, several researchers have found a logarithmic relationship between the legibility of a character and its luminance (Smyth 1947; Forbes, Saari, Greenwood, Goldblatt, and Hill, 1976). Consequently, the logarithm of the combined group visibility density data was taken and plotted against response time using a linear regression model, resulting in the following equation:

$$RT\ (Sec) = -2.996/Log(V) + 1.527, \quad r^2 = 0.831 \qquad (3)$$

A graph of this equation can be found in Figure 72. As can be seen from this figure, a better fit to this data would result using a quadratic equation. The quadratic counterpart to equation 3 is:

$$RT\ (Sec)=(4.39/Log(V^2))-(3.793/Log(V))+1.415, \quad r^2=0.936. \qquad (4)$$

A graph of the quadratic equation can be found in Figure 73. The better fit resulting from the quadratic regression model is due to the reaction time floor effect discussed previously.

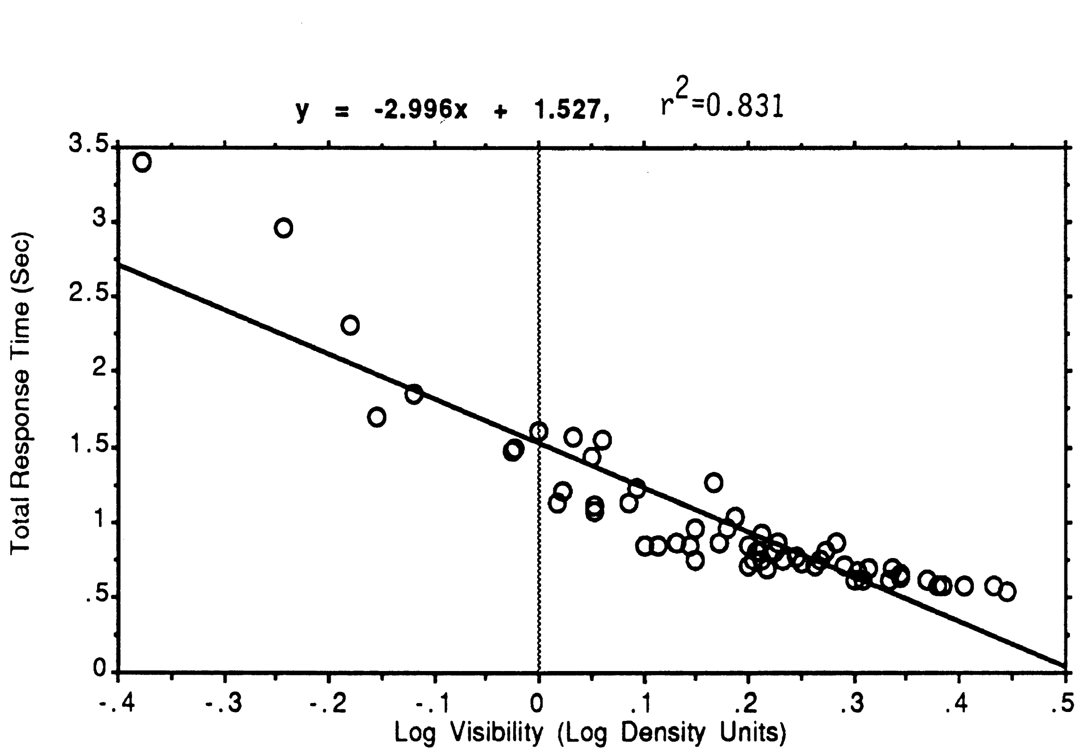


Figure 72. Response Time for All Participants As a Function of Log Visibility (Linear Model).

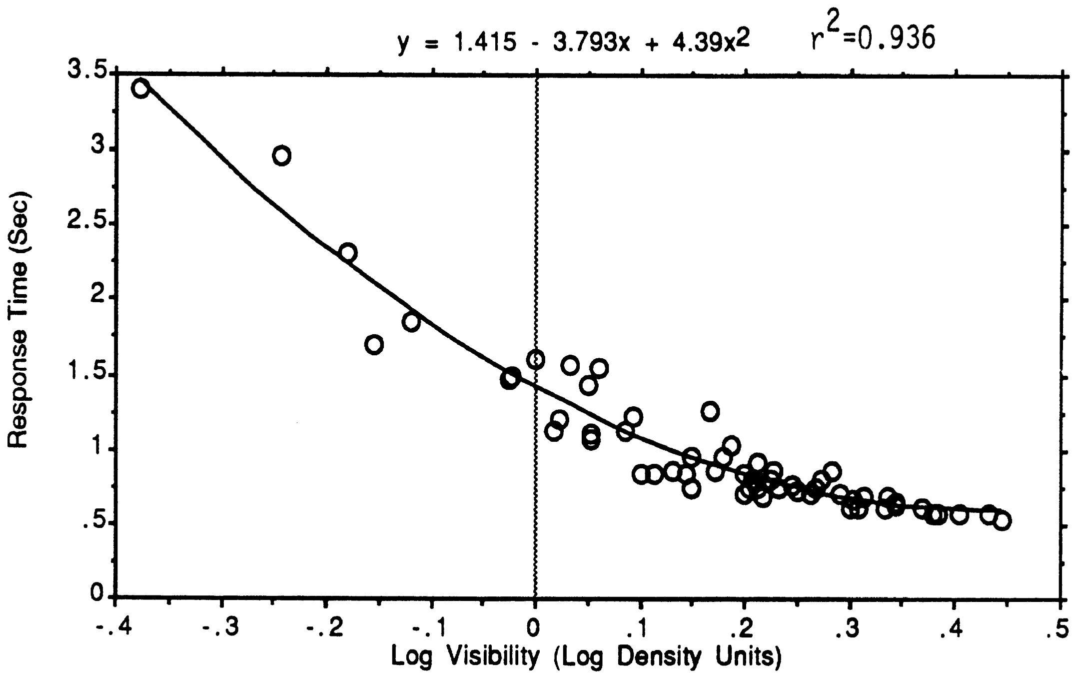


Figure 73. Response Time for All Participants As a Function of Log Visibility (Quadratic Model).

Using the data from the three older subject groups, Sauter and Kerchaert propose a relation to predict a letter's visibility for a typical older driver as a function of illumination as:

Visibility (Density Units)=V=0.47*log(I)-(1.12 * A)+1.38

Where:

I = Target Illumination (fc)
A = Visual Acuity (inverse minutes of arc)

A plot of this function can be found in Figure 74.

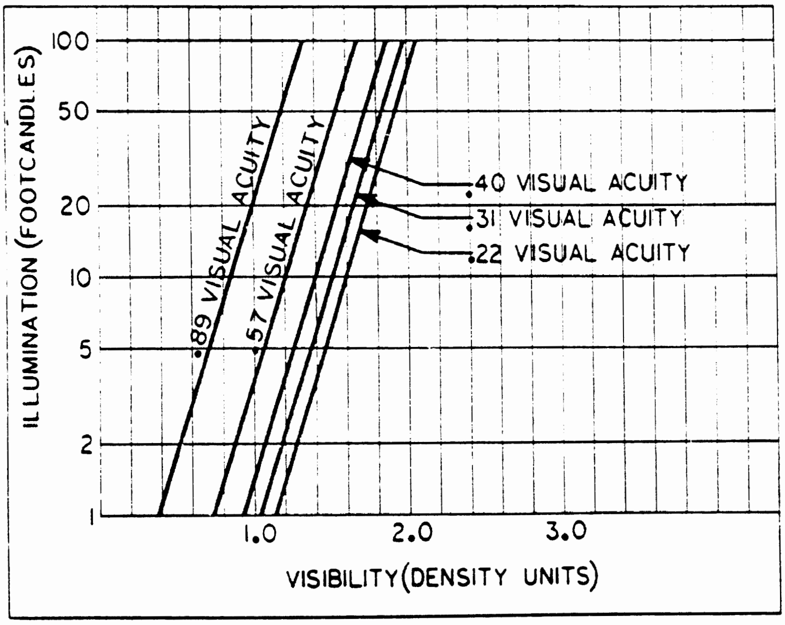


Figure 74. Visibility As a Function of Illumination of Specific Alphanumeric Groupings.

Thus, this study proposes a method for predicting response times for older drivers using data from people of any age or visual acuity. First, several visibility measurements are taken for one of the targets used in this experiment for various illumination levels. These data are plotted in Figure 74. The algebraic difference in visibility density units between the person's data and the data of Figure 74 becomes the person's individual correction factor. For any future visibility readings from this person, the corresponding visibility for a typical older driver would be the person's visibility measurement less the correction factor. Using this corrected visibility measurement, such factors as response time and perception time for the older person can be predicted using Figures 69 - 74.

The advantage of this method is that it removes the need to recruit a large sample of older drivers for studies on visibility, which may be both costly and time-consuming. The primary disadvantages are that the procedure assumes that only perception affects visibility (i.e., that the effect of visual search on visibility can be ignored), and that changes in perception are controlled only by illumination changes. Further, estimates can be made only after a display has been designed, not predicted beforehand.

Mourant and Langolf (1976)

This article presents luminance specifications for automobile instrument panels to provide satisfactory visibility for drivers over 45 years. Three groups of people were tested -- 10 drivers over 45 who normally wore glasses to drive, 10 who did not, and 3 younger drivers who did not wear glasses.

On each trial of the experiment the name of one of seven instrument panel labels (e.g., fan) was read aloud while the driver fixated at a light far away. The light turned off, the driver shifted their gaze to a screen 81.3 cm (32.0 inches) away, and a slide showing four labels of equal length was displayed. The driver then pressed a button corresponding to the target label called out. To provide an approximate 95% luminance threshold criteria, testing continued until a person's luminance threshold (the last point at which a driver could respond correctly in at least 9 out of 10 trials) was identified.

Four letter sizes (strokewidths) ranging from 0.23 to 0.84 cm (e.g., 1.6 to 6.0 minutes of arc) were used. The height to strokewidth ratio was 6:1. For each letter size, the letter:background luminance contrast ratios tested were 1.25:1, 2:1, and 25:1. For each letter size/contrast combination, luminance levels available were in the range from 0.34 to 68.52 cd/m^2.

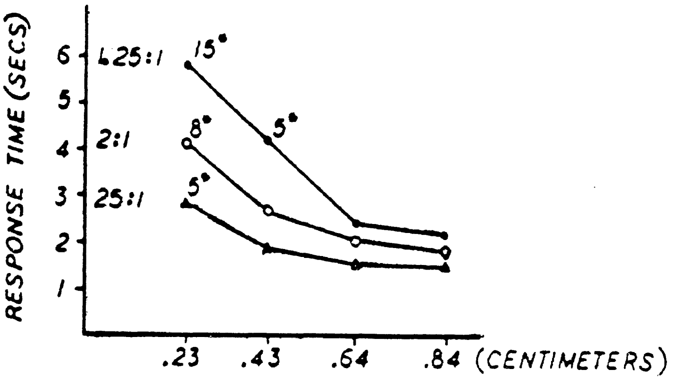
Shown in Table 21 are the luminance levels needed to make the letters just barely readable to 95% of the drivers over 45 years of age. Luminance thresholds could not be supplied for the first four conditions because not all participants could see under these conditions regardless of letter luminance. Further, it should be noted that many of the luminance threshold levels in this table could produce discomfort or disability glare for some drivers, especially for older drivers. Specifically, luminances greater than 3.426 to 5.139 cd/m^2 (1.0 to 1.5 fL) should be avoided, and consequently, the only viable recommendations are those underlined in Table 21. Therefore, the best way to achieve visible displays for older drivers may be to increase letter height rather than increasing luminance levels, although instrument panel size constraints then become a factor.

Table 21. 95% Upper Confidence Limit Minimum Luminance Thresholds.

Letter Size		95% Upper Limit Threshold (cd/m^2)			
StrokeWidth (min)	Height (cm)	Contrast Ratio	Older Glasses	Older No Glasses	Pooled
1.6	0.23	1.25:1	-	-	-
1.6	0.23	2:1	-	-	-
1.6	0.23	25:1	-	-	-
3.0	0.43	1.25:1	-	-	-
3.0	0.43	2:1	18.2	27.8	22.3
3.0	0.43	25:1	3.1	6.8	5.1
4.5	0.64	1.25:1	14.4	48.0	33.6
4.5	0.64	2:1	6.2	6.5	6.2
4.5	0.64	25:1	1.7	1.7	1.7
6.0	0.84	1.25:1	3.4	8.6	6.5
6.0	0.84	2:1	2.1	6.5	4.4
6.0	0.84	25:1	0.3	1.7	1.4

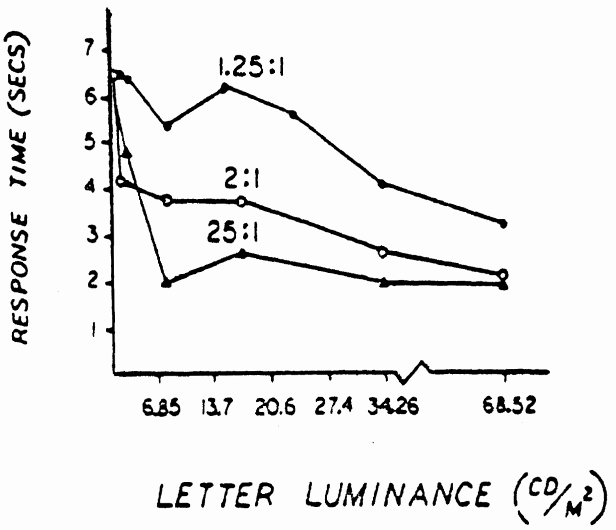
Source: Mourant and Langolf (1976)

A second consideration for legible displays in automobiles is response time, which should be minimized to avoid hazardous driving situations. Shown in Figures 75 and 76 are mean response times as a function of both letter size and letter luminance. As can be seen in these figures, response times vary between 1.5 and 5.7 seconds. (For a car traveling 65 miles per hour, these response times correspond to distances traveled between 143 and 543 feet.)



Source: Mourant and Langolf (1976)

Figure 75. Effect of Letter Size and Contrast Ratio on Response Time.



Source: Mourant and Langolf (1976)

Figure 76. Relationship Between Letter Luminance and Response Time for Three Contrast Ratios.

Since the Mourant and Langolf paper did not include a regression analysis of the response time data (nor did the original technical report) and the model is particularly important, further analysis of the data was conducted.

A stepwise forward regression analysis was carried out with size, contrast ratio, inverse contrast ratio, character luminance, and the inverses and logarithms of those factors as terms from which to select. The product of contrast ratio and size was also included because that interaction is often significant. These choices were made based upon the literature described elsewhere in this report.

Using .05 as the level for inclusion and .1 for exclusion the following equation resulted:

$$\text{Response Time (seconds)} = RT = 5.82 - 13.03H - .70\text{Log}(L) + 2.94/C$$

where:

H = Height (inches)
L = Character Luminance (foot-Lamberts)
C = Contrast Ratio

This equation accounted for almost 65% of the variance in the response times.

Summary

Both studies presented in this section attempted to measure the legibility of characters on instrument panels for older drivers (over 45 years old). Characters were achromatic. The Sauter and Kerchaert study developed a method to predict visual capabilities of older drivers by adjusting the data from younger drivers obtained under similar conditions. The Mourant and Langolf study developed specific recommendations for instrument panel character luminance as a function of character height and contrast ratio.

Both studies employed similar tasks and age groups for collecting their data. Specifically, they had people over 45 years old fixate on a target in front of them and, when cued, had them look at a simulated instrument panel, recognize a word or letter, and respond. Further, both studies were concerned with threshold visibility measurement. Sauter and Kerchaert found threshold visibility through the use of a Luckiesh-Moss visibility meter, while Mourant and Langolf found threshold luminance levels necessary for 95% correct stimulus identification.

For luminance contrast ratios of 25:1, Mourant and Langolf recommend luminance levels of 6.8 cd/m2 for characters 0.43 cm (.17 in) high, and 1.7 cd/m2 for both 0.64 (.25 in) and 0.84 cm (.33 in) characters. These recommendations should be used with some care since they emphasize what an older driver can just barely see, not what they can see easily.

KEY FINDINGS

This report deals with several key questions:

1. *How do fundamental lighting variables (luminance contrast and illumination level) and exposure duration affect people's ability to detect simple visual targets?*
2. *What is the effect of chromatic contrast on legibility?*
3. *What is the effect of font on the legibility of text?*
4. *How well do adults see?*
5. *What expressions are there to predict the legibility of text for various applications?*

These questions are examined one at a time here.

1. How do fundamental lighting variables (luminance contrast and illumination level) and exposure duration affect people's ability to detect simple visual targets?

The most commonly referred to summary of the effect of lighting variables on legibility thresholds are the **Cobb and Moss curves.** Shown below are their data as replotted by Luckiesh and Moss (1937). In brief, these curves show that for fixed visual angles, the primary factor affecting the threshold is the contrast ratio, with illumination levels and exposure duration having a secondary effect. In general, legibility thresholds are fairly independent of distance.

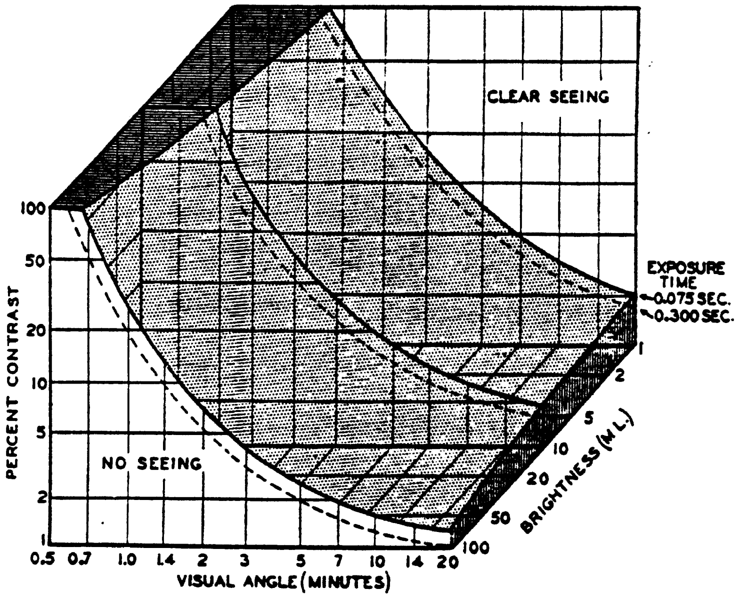


Figure 77. Luckiesh and Moss (1937) Version of Cobb and Moss Curves.

A number of efforts have been made to express threshold functions as equations. Based on the work of **Moon and Spencer (1944)** the following emerged.

For Landolt rings the legibility threshold is:

Minimum Perceptible Angle (radians) =

$$V = (118.3x10^{-6} / Lb) * (1.433 + Lb^{.5})^{2}$$

where:

Lb = Background Luminance (millilamberts)

Contrast ratios are not described and assumed to be high.

For single bar objects the equation for V is:

$$V = (1.950*10^{-6} / Lb) * (1.433 + Lb^{.5})$$

Also, for Landholt rings where the percent correct response is unknown legibility thresholds can be estimated as follows:

$$V=(118.3x10^{-6}/ Lb) * ((y-1/8)/(1-7))^{.05} * (0.412 + B^{.33})^{3}$$

An alternative (and more complex) predictor of the legibility threshold for Landholt C's is by Farber (based on **Blackwell's** data). It is as follows:

Contrast Threshold =
$$Cth = Cx * (0.923/n) * [(S/t*Lb)^{.4} + 1]^{2.5})$$

where:

Cth = Threshold Contrast
d = Target Diameter (minutes of arc)
Cx = Target Size Factor
if $d<=10$, $Cx = 3*(.37)^{\log2(d)}$
if $d>10$, $Cx = .106 - .0006d$
Lb = Background Luminance (cd/m2)

$n = ((S/100t)^{.4} + 1)^{2.5}$
$S = 0.5900 - 0.6235\log d - s$
(effect of age on slope of RCS function of luminance)

s = Adjustment Parameter
if Age 20-44, s= 0
Age 44-64, s=.00406 (A-44)
Age 64-80, s=.0812 + .00667 (A-64)

t = Relative Equivalent Occular Transmittance
(loss due to age)
if Age 20-30, $\log t=0$
Age 30-44, $\log t=.01053\,(A-30)$
Age 44-64, $\log t=.1474 - .0134\,(A - 44)$
Age 64-80, $\log t=.4154 - .0175\,(A - 64)$

The critical group of terms in the model is the Relative Contrast Threshold (RCS), and expression used to adjust to fit a basic detection function to a wide variety of circumstances. It is expressed as:

$$RCS = n * [(S/tL)^{.4} + 1]^{-2.5}$$

Finally, to adjust for the change in contrast threshold with age, Farber describes a multiplier used by Blackwell as follows:

if Age 20-42, $m1 = 1.000 + .00795\,(A - 20)$
Age 42-64, $m1 = 1.175 + .0289\,(A - 42)$
Age 64-80, $m1 = 1.811 + .1873\,(A - 64)$

Hence, the literature contains a rather complete description of the effects of lighting variables on performance both in the form of figures (the Cobb and Moss Curves) and equations. Of particular note is the work of Blackwell and readers are encouraged to examine it further.

2. What is the effect of chromatic contrast on legibility?

In general, the effects of chromatic contrast on legibility are relatively small. Two expressions of chromatic contrast appear in the literature, one based on the **CIELAB** data (CIE Yu'v') and the other based on **CIELUV**. They are listed below in that order.

$$\text{Delta\_E} = [\ (\text{Delta\_L}^*)^2 + (\text{Delta\_u}^*)^2 + (\text{Delta\_v}^*)^2\ ]^{1/2}$$

where:

$L^* = 116(Y/Y0)^{1/3} - 16,\ Y/Y0>.01$
$u^* = 13L^*(u'-u0')$
$v^* = 13L^*(v'-v0')$
$u' = 4X/(X + 15Y + 3Z)$
$v' = 9Y/(X + 15Y + 3Z)$

$$\text{Delta\_E}(Yu'v') = [(155\ TB/M)^2 + (367\ U)^2 + (167\ V)^2]^{0.5}$$

where TB = difference in luminance between text and background
M = maximum luminance of text or background
U = difference between u' coordinates of text and background
V = difference between v' coordinates of text and background

The ANSI standard for office workstations requires that characters presented on screens have Delta_E values in excess of 100 (CIE Lu*v*).

3. What is the effect of font on the legibility of text?

In general, the literature shows that the effects of varying font on performance are relatively small. However, that does not mean that font should be ignored, for font modifications are quite straightforward.

Beyond that general conclusion, several key points emerge. First, for seven-segment displays confusion errors are very predictable. The frequency of particular confusions depends upon the number of segments in which the character pair in question differ. For small sets (e.g., numbers only) Van Nes and Bouma offer design suggestions to minimize reading errors.

With regard to dot matrix characters, the key studies are those of **Snyder and Maddox** (**1978**) and **Shurtleff's book.** They show that characters should be at least 7x9 for easy reading with 9x11 preferred. Dots should be close together and round.

4. How well do adults see?

To design a product for people to use, engineers need to know what user capabilities are. In the case of displays the important question is, "How well do people see?"

There are two particularly useful studies of how well adults see. A British study focuses particularly on drivers (**Davison and Irving, 1980**) while the U.S. work (**Roberts, 1964**), carried out as part of the national health survey, concerns all adults. The U.S. data follow (Table 22). Note the large number of drivers whose corrected acuity is 20/40 or worse, especially in the older groups.

5. What expressions are there to predict the legibility of text for various applications?

There are several formulas in the literature that summarize experimental work on the legibility of text on displays. The relationship proposed by **Peters and Adams** (**1959**) is often used by human factors engineers because it appeared in a commonly referred-to edition of a human factors textbook. Unfortunately, their paper does not describe the empiric basis for the equation so it should be used with caution. According to them, required letter height can be estimated as follows:

Table 22. Percentage of Adults Reaching Specified Acuity Levels - Corrected Far Acuity.

Sex and acuity level	Total, 18-79 years	18-24 years	25-34 years	35-44 years	45-54 years	55-64 years	65-74 years	75-79 years
Both sexes	Percent distribution							
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
20/10 or better	1.5	2.4	2.6	1.7	0.9	0.5	0.2	-
20/15	38.4	57.8	57.7	52.8	31.1	11.2	4.4	1.6
20/20	32.9	27.5	29.3	35.5	39.0	38.4	27.6	13.4
20/30	17.7	10.2	7.4	7.1	20.5	33.5	38.9	38.6
20/40	4.5	1.4	1.5	1.4	4.2	8.5	12.6	18.0
20/50	1.8	0.4	0.3	0.8	1.5	2.6	6.2	9.3
20/70	0.9	-	0.4	0.3	0.5	1.6	2.0	7.9
20/100	1.5	0.0	0.5	0.2	1.6	2.5	5.3	7.9
20/200	0.4	0.2	0.2	0.1	0.4	0.7	0.9	1.3
Less than 20/200	0.4	0.1	0.1	0.1	0.3	0.5	1.9	2.0
Men								
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
20/10 or better	1.9	2.9	4.0	2.0	1.4	0.2	-	-
20/15	42.8	61.3	63.5	58.4	37.3	12.2	4.5	3.2
20/20	31.0	25.4	25.0	30.9	36.6	39.1	32.4	15.7
20/30	15.8	8.6	5.4	5.8	18.0	32.5	35.0	38.0
20/40	4.2	1.6	1.2	1.4	3.2	8.8	11.9	19.8
20/50	1.7	0.2	0.2	0.5	1.9	2.3	5.8	10.3
20/70	0.7	-	0.1	0.5	0.3	1.4	2.8	3.2
20/100	1.4	-	0.4	0.1	0.9	2.8	5.9	8.3
20/200	0.3	-	0.2	0.2	0.4	0.2	1.2	-
Less than 20/200	0.2	-	-	0.2	-	0.5	0.5	1.5
Women								
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
20/10 or better	1.1	1.9	1.3	1.5	0.5	0.8	0.4	-
20/15	34.7	54.9	52.6	47.5	25.3	10.3	4.3	-
20/20	34.6	29.2	33.2	39.8	41.1	37.8	23.8	11.2
20/30	19.5	11.6	9.3	8.2	22.9	34.4	42.1	39.1
20/40	4.7	1.2	1.7	1.5	5.2	8.3	13.1	16.4
20/50	1.8	0.6	0.3	1.0	1.1	2.8	6.4	8.4
20/70	1.0	-	0.7	0.1	0.7	1.8	1.4	12.5
20/100	1.6	0.1	0.6	0.3	2.3	2.2	4.8	7.4
20/200	0.4	0.3	0.2	-	0.3	1.2	0.7	2.6
Less than 20/200	0.6	0.2	0.1	0.1	0.6	0.4	3.0	2.4

Letter Height (inches) = $H = .0022D + K1 + K2$

where:

D = Viewing Distance (inches)
K1 = correction factor for illumination and reading situation
= 0.06 for illumination > 1.0 fc, favorable reading conditions
= 0.16 for illumination > 1.0 fc, unfavorable conditions or illumination < 1.0 fc, favorable conditions
= 0.26 for illumination < 1.0 fc, unfavorable conditions
K2 = Correction for importance
= 0.075 for emergency labels, counters, scales, legend lights
= 0.0 for other (unimportant) panel markings

Sauter and Kerchaert (1972) took a different approach by trying to develop a legibility model for one group of people and then trying to adapt it to others using scores from a Luckiesh-Moss Visibility Meter. Their prediction for response time to look for a target ahead and identify a character in one of three locations on an instrument panel is as follows:

Response Time (sec) = $RT = (1.58/V^2) - (0.55/V) + 0.56$

where:

V = visibility meter reading in density units

Based on the data provided in their paper, V, the meter reading, can be predicted as follows:

Visibility (density units)=$V=0.47*\log(I)-(1.12 * A)+1.38$

where:

I = Target Illumination (fc)
A = Visual Acuity (inverse minutes of arc)

It should be noted that these are not the only ways in which visibility or response time can be expressed. Readers should examine the "Application Independent Studies of Legibility" section for alternatives.

The work that is closest in approach to the problem of interest here is that of **Mourant and Langolf (1976)**. They had people perform a task that is reasonably close to that of driving. Their data was reanalyzed in this review. Based on that data, response time can be predicted as follows:

Response Time (seconds) = $RT = 5.82 - 13.03H - .70\log(L) + 2.94/C$

where:

H = Height (inches)
L = Character Luminance (foot-Lamberts)
C = Contrast Ratio

A study that is specifically concerned with LED and LCD displays is that of **Duncan and Konz (1976).** It too should be applied with some caution to automotive applications, as their viewing distances were far greater than the standard instrument panel viewing distance. For instrument panels, near acuity, not far acuity, is of interest. A rearrangement of their data shows that character heights can be estimated as follows:

$$\text{Height (cm)} = H = .0015De + .0519(H{:}Sw) - .3499$$

where:

De = No Error Viewing Distance (cm)
H:Sw = Height:Strokewidth Ratio.

\- or -

$$H = .0038Dp + .0385(H{:}Sw) - .0864$$

where:

De = Preferred Viewing Distance (cm)
H:Sw = Height:Strokewidth Ratio.

Notice that the preferred character size is about double the character size required for good performance (no errors). Other studies have suggested even larger ratios.

A fairly commonly used rule of thumb (the Bond Rule) was developed by **Smith (1979).** His simple rule ignores lighting variables entirely and assumes the only important factor is the visual angle of the character. This procedure overestimates the character size required in many cases. However, it has the advantage of being a relationship that is easily memorized and is often used where time for a detailed analysis is not available. According to the Bond Rule, the required character height is as follows:

Height = .007 x Viewing Distance (D, same units as height)

The most commonly specified design requirements are in the de facto human engineering standard, **Military Standard 1472C (U.S. Department of Defense, 1981b).** They require that characters be in a range of sizes. Readers should bear in mind the standard presents minimum values, not desired ones. (See Table 23.)

Table 23. Label Size Versus Luminance.

Markings	Height*	
	$3.5\ cd/m^2$ (1 fL) or BELOW	ABOVE $3.5\ cd/m^2$ (1 fL)
For critical markings, with position **variable** (e.g., numbers on counters and settable or moving scales)	5-8 mm (0.20-0.31 in)	3-5 mm (0.12-0.20 in)
For critical markings, with position **fixed** (e.g., numerals on fixed scales, controls, switch markings, or emergency instructions)	4-8 mm (0.16-0.31 in)	2.5-5 mm (0.10-0.20 in)
For **non-critical** markings (e.g., identification labels, routine instructions, or those required	1.3-5 mm (0.05-0.20 in)	1.3-5 mm (0.05-0.20 in)

Source: U.S. Department of Defense (1981b)

A more detailed procedure was proposed by **Howett (1983)**. This procedure, commonly called the NBS method, is based upon visual acuity data for adults. Unlike some of the expressions given previously, it includes the viewer visual acuity as a parameter. There are three steps to the procedure.

One, compute the luminance contrast.

$$\text{Contrast } (\%) = C = ((Lb - Lt) / Lb) * 100 \quad (\text{assumes } Lb>Lt)$$

where: Lb = Background Luminance
Lt = Target Luminance

Since C is dimensionless, the units of Lb and Lt do not matter here as long as they are the same.

Two, compute the relative Snellen Acuity, S. The Snellen acuity is a measure of how well people see. It was described earlier. In this instance, the acuity of interest is that of the worst case viewer.

$$S = Sd * (85 / Lb)^{.213} * (90 / C)^{.532}$$

where: Sd = Denominator in the Snellen ratio.
(If a viewer has 20/40 visual acuity, use 40.)
Lb = Background Luminance (cd/m2)

Three, compute the Character Height, H (m).

$$H = (H{:}Sw) * 1.45 * 10^{-5} * S * D$$

where: H:Sw = Height to Strokewidth Ratio (for 6:1 use 6)
D = Viewing Distance (m)

Payne's 1983 work was concerned with a specific display technology, LCD. The advantage of his work over the Howett study just mentioned is that it allows one to relate performance to character size as opposed to giving a single number, the required minimum size. According to Payne the error rate in reading groups of numbers can be estimated as follows:

$$\text{Error Rate } (\%) = E = 1.52 + .02Bl - 1.40Ca + .02Va - .0006Ea$$

where:

Bl = Back Light Luminance (0 to 122 cd/m^2)
Ca = Character Subtense Angle (0.025 to 1.34 degrees)
Va = Viewing Angle (0 to 60 degrees)
Ea = Ambient Light Illumination (20 to 1500 lx).

A model for legibility that seems to be particularly well suited to conditions of high illumination levels was developed by **Rogers, Spiker, and Cicinelli (1986).** Unfortunately their analysis is incomplete, expressing response time using three functions, not one. They are shown below.

$RT = 100(CR^{-1})$ for 3.4 cd/m^2 (1 fL)

$RT = 31(CR^{-2})$ for 34.3 cd/m^2 (10 fL)

$RT = 8(CR^{-2})$ for 257 cd/m^2 (75 fL)

where: RT = Response Time
CR = Contrast Ratio

A very recent addition to the legibility literature is the work of **Sawyer and Talley (1987).** They developed a table showing the visual angle resulting from various viewing distance-character height combinations along with recommendations for size. As noted earlier, their recommendations are based on their assessment of the literature and not on any particular study. Their larger sizes tend to be close to the values recommended by the Bond Rule. A table showing their recommendations follows. The table has proven to be popular with designers because it is reasonably straightforward.

Summary

Clearly, there are many formulas in the vision and human factors literature that predict character size as a function of lighting parameters, visual acuity, and other factors. Most of them concern static viewing conditions where no visual search is

Viewing Distance in inches (centimeters)

Character Height in. (cm)	18.0 (45.7)	20.0 (50.8)	22.0 (55.9)	24.0 (61.0)	26.0 (66.0)	28.0 (71.1)	30.0 (76.2)	32.0 (81.3)	34.0 (86.4)	36.0 (91.4)
0.04 (0.10)	7.6	6.9	6.3	5.7	5.3	4.9	4.6	4.3	4.0	3.8
0.06 (0.15)	11.5	10.3	9.4	8.6	7.9	7.4	6.9	6.4	6.1	5.7
0.08 (0.20)	15.3	13.8	12.5	11.5	10.6	9.8	9.2	8.6	8.1	7.6
0.10 (0.25)	19.1	17.2	15.6	14.3	13.2	12.3	11.5	10.7	10.1	9.6
0.12 (0.30)	22.9	20.6	18.8	17.2	15.9	14.7	13.8	12.9	12.1	11.5
0.14 (0.36)	26.7	24.1	21.9	20.1	18.5	17.2	16.0	15.0	14.2	13.4
0.16 (0.41)	30.6	27.5	25.0	22.9	21.2	19.6	18.3	17.2	16.2	15.3
0.18 (0.46)	34.4	30.9	28.1	25.8	23.8	22.1	20.6	19.3	18.2	17.2
0.20 (0.51)	38.2	34.4	31.3	28.7	26.4	24.6	22.9	21.5	20.2	19.1
0.22 (0.56)	42.0	37.8	34.4	31.5	29.1	27.0	25.2	23.6	22.2	21.0
0.24 (0.61)	45.8	41.3	37.5	34.4	31.7	29.5	27.5	25.8	24.3	22.9
0.26 (0.66)	49.7	44.7	40.6	37.2	34.4	31.9	29.8	27.9	26.3	24.8
0.28 (0.71)	53.5	48.1	43.8	40.1	37.0	34.4	32.1	30.1	28.3	26.7
0.30 (0.76)	57.3	51.6	46.9	43.0	39.7	36.8	34.4	32.2	30.3	28.7
0.32 (0.81)	61.1	55.0	50.0	45.8	42.3	39.3	36.7	34.4	32.4	30.6

Visual Angle in minutes of arc

Underlined area is the preferred range of visual angle.
Tinted area is the minimum recommended range.

required, and color is not usually a factor. For automotive applications search is clearly an important factor. Further, many of them were based on test conditions involving far acuity, not near acuity that is critical to reading instrument panels. While viewing distance is clearly a secondary factor, it should nonetheless be included (or fixed) in a model that predicts the legibility of instrument panel displays. Also, the trend is for studies to emphasize mimimum legibility. When designing a vehicle, the goal is to make the displays easy to read, not barely readable. The one exception to these criticisms is the work of Mourant and Langolf (1976). Their data has been reanalyzed, and a regression model that predicts response time appears earlier in this section.

What expression should engineers use to predict required character sizes on instrument panels? The authors would argue for employing the model developed in this report based on the Mourant and Langolf data for a start, and then adjusting the figures to take into account the result obtained using Howett's expression. If a quick answer is needed, use the Bond Rule.

Because of these shortcomings of the literature, a series of experiments were conducted to determine how instrument panel displays should be evaluated (Bos, Green, and Kerst, 1988) and how big numbers on numeric speedometers should be so they can be easily read (Boreczky, Green, Bos, and Kerst, 1988). This research is being carried out in parallel with the writing of this review.

- Key Findings -

REFERENCES

Adams, S., Goldstein, S., Zeltner, K., Ratanaproeksa, P., and Green, P. (1988). Legibility Abstracts from the UMTRI Library (Technical Report UMTRI-88-4). Ann Arbor, MI: The University of Michigan Transportation Research Institute, February.

Beamon, W.S. and Snyder, H.L. (1979). An Experimental Evaluation of the Spot Wobble Method of Suppressing Raster Structure Visibility (Report Number AMRL-TR-75-63). Wright-Patterson Air Force Base, OH: U.S. Air Force Aerospace Medical Research Laboratory.

Bendix Radio Division (1959). Design for Legibility of Visual Displays (Report Number 481-1016-97A). Baltimore, MD: Bendix Aviation Corporation.

Billmeyer, F.W. Jr. and Saltzman, M. (1981). Principles of Color Technology. New York: Wiley.

Blackwell, H. (1946). Contrast Thresholds of the Human Eye. Journal of the Optical Society of America, November, 36, 624.

Blackwell, H. (1952). Brightness Discrimination Data for the Specification of Quantity of Illumination. Illuminating Engineering, November, XLVII(11), 602-609.

Blackwell, H. (1955). Use of Performance Data to Specify Quantity and Quality of Interior Illumination. Illuminating Engineering, June, L(6), 286-299.

Blackwell, H.R. (1981). Description of a Comprehensive Family of Relative Contrast Sensitivity (RCS) Functions of Luminance to Reflect Differences in Size of Task Detail, Task Eccentricity and Observer Age, Journal of the Illuminating Engineering Society, October, 52-63.

Blackwell, H.R. (1982). A Comprehensive Quantitative Method for Prediction of Visual Performance Potential as a Function of Reference Luminance. Journal of the Illuminating Engineering Society, October, 40-63.

Blackwell, O.M. and Blackwell, H.R. (1980a). Individual Responses to Lighting Parameters for a Population of 235 Observers of Varying Ages, Journal of the Illuminating Engineering Society, July, 205-232.

Blackwell, H.R. and Blackwell, O.M. (1980b). Population Data for 140 Normal 20-30 Year Olds for Use in Assessing Some Effects of Lighting Upon Visual Performance, Journal of the Illuminating Engineering Society, April, 9(3), 158-174.

- *References* -

Boff, K.R., Kaufman, L., and Thomas, J.P. (1986). Handbook of Perception and Human Performance (Volumes I and II). New York: Wiley.

Boreczky, J.S., Green, P., Bos, T.L., and Kerst, J.O. (1988). Effects of Contrast, Illumination, and Color on the Legibility of Numeric Speedometers (Technical Report UMTRI-88-36). Ann Arbor, MI: The University of Michigan Transportation Research Institute.

Boynton, R.M., Nagy, A.L., and Olson, C.X. (1983). A Flaw in Equations for Predicting Chromatic Differences. Color Research and Applications, Summer, 8(2), 69-74.

Brown, F. (1951). A Study of the Requirements for Letters, Numbers, and Markings to be Used on Trans-Illuminated Aircraft Control Panels (Part 3 - The Effect of Stroke Width and Form upon the Legibility of Numerals), (Report TED no. NAMEL-609-3). Philadelphia, PA: Naval Air Material Center.

Brown, F. (1953). A Study of the Requirements for Letters, Numbers, and Markings to be Used on Trans-Illuminated Aircraft Control Panels (Part 4 - Legibility of Uniform Stroke Capital Letters as Determined by Size and Height to Width Ratio and as Compared to Garamond Bold) (Report TED no. NAMEL-609-4). Philadelphia, PA: Naval Air Material Center.

Carlson, C.R. and Cohen, R.W. (1978). Visibility of Displayed Information: Image Descriptors for Displays (Report Number ONR-CR213-120-4F). Arlington, VA: Office of Naval Research.

Carter, R.C. and Carter, E.C. (1983). CIE L*u*v* Color Difference Equations for Self-Luminous Displays. Color Research and Application, Winter, 8(3), 252-253.

Chapanis, A. (1949). How We See: A Summary of Basic Principles (Chapter 1) in National Research Council, Human Factors in Undersea Warfare. Washington, D.C.: National Research Council, Committee on Undersea Warfare.

Chapanis, A., Garner, W.R., and Morgan, C.T. (1949). Applied Experimental Psychology: Human Factors in Engineering Design. New York: John Wiley & Sons.

Charman, W.N. and Olin, A. (1965). Tutorial: Image Quality Criteria for Aerial Camera Systems. Photographic Science and Engineering, 9, 385-397.

Christ, R. (1975). Review and Analysis of Color Coding Research for Visual Displays. Human Factors, December, 17(6), 542-570.

Clear, R. and Berman, S. (1983) A New Look at Models of Visual Performance. Journal of the Illuminating Engineering Society, July, 242-250.

Cobb, P. and Moss, F. (1928). The Four Variables of the Visual Task Threshold. Journal of the Franklin Institute, June, 205(1230), 831-847.

Commission Internationale de l'Eclairage (International Commission on Illumination) (1972). A Unified Framework of Methods for Evaluating Visual Performance Aspects of Lighting (CIE report 19). Paris: Commission Internationale de l'Eclairage.

Cornog, D.Y. and Rose, F.C. (1967). Legibility of Alphanumeric Characters and Other Symbols II. A Reference Handbook (Miscellaneous Publication 262-2). Washington, D.C.: National Bureau of Standards.

Davison, P.A. and Irving, A. (1980). Survey of Visual Acuity of Drivers (TRRL Laboratory report 945). Crowthorne, Berkshire, U.K.: Transport and Road Research Laboratory, Road User Characteristics Division.

Dees, J.W. (1964). Light and Human Visual Characteristics (Report A240). St. Louis, MO: McDonnell Aircraft Corporation.

Duncan, J. and Konz, S. (1976). Legibility of LED and Liquid-Crystal Displays. Proceedings of the Society for Information Display, 17(4), 180-186. Farnborough, Hants, UK: Royal Aircraft Establishment.

Ellis, N. and Hill, S. (1978). A Comparative Study of Seven-Segment Numerics. Human Factors, December, 20(6), 655-660.

Farber, E. (1988). Revising the DETECT Seeing Distance Model (SAE Paper Number 880716). Warrendale, PA: Society of Automotive Engineers.

Forbes, T.W., Saari, B.B., Greenwood, W.H., Goldblatt, J.G., and Hill, T.E. (1976). Luminance and Contrast Requirements for Legibility and Visibility of Highway Signs. Transportation Research Record, January, (562), 59-72.

Fortuin, G. (1970). Lighting: Physiological and Psychological Aspects. Ergonomics and Physical Factors, 237-259. Geneva, Switzerland: International Labor Office.

Foster, J. (1980). Legibility Research 1972-1980: A Summary. Manchester, U.K.: Manchester Polytechnic, Royal College of Art, Graphic Information Research Unit.

Gallagher, V. and McCunney, W. (1977). Toward the Quantification of Visibility (HEW Contract Number 210-76-0116). Cincinnati, OH: National Institute for Occupational Safety and Health.

Gibney, T.K. (1967). Legibility of Segmented Versus Standard Numerals: A Review (Technical Report AMRL-TR-67-116). Wright-Patterson Air Force Base, OH: U.S. Air Force Aerospace Medical Research Laboratories.

Green, P. (1988). Human Factors and Gauge Design: A Literature Review (Technical Report UMTRI-88-37). Ann Arbor, MI: The University of Michigan Transportation Research Institute, September.

Hecht, S. (1935). The Theory of Visual Intensity Discrimination. Journal of General Physiology, May, (18), 767-789.

Hind, P.R., Tritt, B.H., and Hoffmann, E.R. (1976). Effects of Level of Illumination, Strokewidth, Visual Angle and Contrast on the Legibility of Numerals of Various Fonts. Proceedings of the Australian Road Research Board, Eighth Conference, 46-55. Parkville, Victoria, Australia: Australian Road Research Board.

Howett, G.L. (1983). Size of Letters Required for Visibility as a Function of Viewing Distance and Observer Visual Acuity (Technical Note NBS 1180). Washington, D.C.: National Bureau of Standards.

The Human Factors Society (1988). American National Standard for Human Factors Engineering of Visual Display Terminal Workstations (ANSI/HFS standard 100-1988). Santa Monica, CA: The Human Factors Society.

Illuminating Engineering Society (1972). IES Lighting Handbook (5th ed.) New York: Illuminating Engineering Society.

Kaiser, P.K. (1988). Color Perception and Measurement (Tutorial T27). Presented at the SPIE-ESD International Symposium on Optical Engineering and Industrial Sensing for Advanced Manufacturing Technologies, Dearborn, MI. Bellingham, WA: Society of Photo-Optical Instrumentation Engineers.

Kaneko, T. (1982). An Alternative ESI System Based on Human Visual Acuity. Journal of the Illuminating Engineering Society, April, 171-177.

Kimura, K., Sugiura, S., Hiroaki, S., and Nagai, Y. (1988). Visibility Requirements for Automobile CRT Displays - Color, Contrast and Luminance. (SAE Paper Number 880218). Warrendale, PA: Society of Automotive Engineers.

Kuehni, R.G. (1982). Advances in Color-Difference Formulas, Color Research and Application, Spring, 7(1), 19-23.

Laycock, J. (1985). The Legibility of Passive Displays. Proceedings of the Society for Information Display, 26(2), 89-93. Farnborough, Hants, UK: Royal Aircraft Establishment.

Linfoot, E.H. (1960). Fourier Methods in Optical Image Evaluation. London, U.K.: Focal Press.

Luckiesh, M. (1944). Light, Vision, and Seeing. New York: Van Nostrand.

Luckiesh, M. and Moss, F. (1937). The Science of Seeing. New York: Van Nostrand.

MacAdam, D.L. (1985). Color Measurement. Berlin: Springer-Verlag.

Maddox, M.E., Burnette, J.T. and Gutmann, J.C. (1977). Font Comparisons for 5x7 Dot Matrix Characters, Human Factors, 19(1), 89-93.

McCormick, E.J. (1970). Human Factors Engineering (3rd ed.). New York: McGraw-Hill.

Moon, P. and Spencer, D. (1944). Visual Data Applied to Lighting Design. Journal of the Optical Society of America, October, 34(10), 605-617.

Mourant, R. and Langolf, G. (1976). Luminance Specifications for Automobile Instrument Panels. Human Factors, February, 18(1), 71-84.

Nakane, Y. and Ito, K. (1978). Study on Standard Visual Acuity Curves for Better Seeing in Lighting Design. Journal of Light and Visual Environment, 2(1), 38-44.

Nason, W. and Bennett, C. (1973). Dials Versus Counters: Effects of Precision on Quantitative Reading. Ergonomics, 16, 749-758.

National Research Council (1949). Human Factors in Undersea Warfare. Washington, D.C.: National Research Council.

Olson, P.L. and Bernstein, A. (1979). The Nighttime Legibility of Highway Signs as a Function of Their Luminance Characteristics. Human Factors, April, 21(2), 145-160.

Olson, P.L., Sivak, M., and Egan, J.C. (1983). Variables Influencing the Nighttime Legibility of Highway Signs (Technical Report Number UMTRI-83-36). Ann Arbor, Michigan: The University of Michigan Transportation Research Institute, July.

Overington, I. (1976). Vision and Acquisition. London. U.K.: Pentech.

Overington, I. (1982). Towards a Complete Model of Photopic Visual Threshold Performance. Optical Engineering, 21, 2-13.

Payne, S. (1983). Readability of Liquid Crystal Displays: A Response Surface. Human Factors, April, 25(2), 185-190.

Peters, G.A. and Adams, B.B. (1959). These 3 Criteria for Readable Panel Markings, May 25, 30(21), 55-57.

Plauth, D.W. (1970). The Readability of Segmented and Conventional Numerals, Human Factors, 12(5), 493-497.

Post, D.L., Costanza, E.B., and Lippert, T.M. (1982). Expressions of Color Contrast as Equivalent Achromatic Contrast. Proceedings of the Human Factors Society 26th Annual Meeting, 581-585.

Pratt, W.K. (1978). Digital Image Processing. New York: Wiley.

Radl-Koethe, H. and Schubert, E. (1972). Comparative Studies of the Legibility of Light Emitting Numerals, SID Journal, May-June, 1(1), 5-10.

Rea, M. (1986). Toward a Model of Visual Performance: Foundations and Data. Journal of the Illuminating Engineering Society, 41-57.

Roberts, J. (1964). Binocular Visual Acuity of Adults, United States, 1960-1962 (technical report series 11, number 3). Washington, D.C.: U.S. Department of Health, Education, and Welfare, Public Health Service, June.

Rogers, S., Spiker, V., and Cicinelli, J. (1986). Luminance and Luminance Contrast Requirements for Legibility of Self-Luminous Displays in Aircraft Cockpits. Applied Ergonomics, 17(4), December, 271-277.

Rosell, F.A. (1971). Analysis of Electro-optical Imaging Sensors (Report Number ADM 105). Pittsburgh, PA: Westinghouse Electric.

Sawyer, D. and Talley, W. (1987). Display Legibility Guidelines: A Design Aid. Information Display, December, 13-16.

Sauter, J. and Kerchaert, R. (1972). Relating Instrument Panel Visibility and Driver Perception Time (SAE Paper Number 720232). Warrendale, PA: Society of Automotive Engineers.

Schmidt, I. and Connolly, P. (1966). Visual Considerations of Man, the Vehicle, and the Highway (SAE Paper Number 660004). Warrendale, PA: Society of Automotive Engineers.

Schade, O.H. (1973). Image Gradation, Graininess, and Sharpness in Television and Motion-Picture Systems. Part III: The Grain Structure of Television Images. Journal of the Society of Motion Picture and Television Engineers, 61, 97-164.

Shannon, C.E. and Weaver, V. (1949). The Mathematical Theory of Communication. Urbana, IL: University of Illinois Press.

Shlaer, S., Smith, E., and Chase, A. (1942). Visual Acuity and Illumination in Different Spectral Regions. Journal of General Physiology, March, 25(4), 553-569.

Shurtleff, D.A. (1980). How to Make Displays Legible. La Mirada, CA: Human Interface Design.

Silverstein, L.D. (1987). Display Color Modeling and Algorithmic Color Selection: Phase II Software Development, Technical Background and Software User's Manual. Phoenix, AZ: Honeywell Inc. Systems and Research Center, October.

Silverstein, L., Lepkowski, J., Carter, R., and Carter, E. (1986). Modeling of Display Color Parameters and Algorithmic Color Selection. Advances in Display Technology VI, 624, Bellingham, WA: Society of Photo-Optical Instrumentation Engineers, 26-35.

Silverstein, L.D. and Merrifield, R.M. (1985). The Development and Evaluation of Color Systems for Airborne Applications, Phase I: Fundamental Visual, Perceptual, and Display System Considerations (Technical Report DOT/FAA/PM-85-19). Washington, D.C.: U.S. Department of Transportation, Federal Aviation Administration, July.

Sivak, M. and Olson, P.L. (1985). Optimal and Minimal Luminance Characteristics for Retroreflective Highway Signs. Transportation Research Record, (1027), 53-57.

Smith, S. (1979). Letter Size and Legibility. Human Factors, December, 21(6), 661-670.

Smith, S.L. (1978). The Limited Readability of Lansdell Numerals, Human Factors, 20(1), 57-64.

Smyth, J.S. (1947). The Brightness and Legibility at Night of Road Traffic Signs. Illuminating Engineering Society Transactions, 12(4), 71-94.

Snyder, H.L. Image Quality: Measures and Visual Performance as cited in Tannas, L. (1985). Flat-Panel Displays and CRT's. New York: Van Nostrand Reinhold, 54-88.

Snyder, H.L. and Maddox, M.E. (1978). Information Transfer from Computer-Generated Dot-Matrix Displays (Technical Report HFL-78-3/ARO-78-1). Research Triangle Park, NC: U.S. Army Research Office.

Snyder, H.L. and 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.

Task, H.L. and Verona, R.W. (1976). A New Measure of Television Display Quality Relatable to Observer Performance (Report Number AMRL-TR-76-73). Wright-Patterson Air Force Base, OH: U.S. Air Force Aerospace Medical Research Laboratory.

Trezona, P. (1983). Luminance Level Conversions to Assist Lighting Engineers to Use Fundamental Visual Data. Lighting Research & Technology, 15(2), 83-88.

Uhlaner, J.E. (1941). The Effect of Thickness of Stroke on the Legibility of Letters. Proceedings of the Iowa Academy of Science, 48, 319-324.

U.S. Department of Defense (1975). Military Standard - Human Factors Engineering Design for Army Materiel (MIL-HDBK-759A). Washington, D.C.: U.S. Department of Defense.

U.S. Department of Defense (1981). Military Standard - Human Engineering Design Criteria for Military Systems, Equipment, and Facilities (MIL-STD-1472C). Washington, D.C.: U.S. Department of Defense.

Van Meeteren, A. (1973). Visual Aspects of Image Intensification. Soesterberg, The Netherlands: Institute for Perception TNO.

Van Nes, F. (1972). Determining Temporal Differences with Analogue and Digital Time Displays. Ergonomics, January, 15(1), 73-79.

Van Nes, F. and Bouma, H. (1980). On the Legibility of Segmented Numerals. Human Factors, August, 22(4), 463-474.

Welsh, K.W., Rasmussen, P.G., and Vaughan, J.A. (1977). Readability of Alphanumeric Characters Having Various Contrast Levels as a Function of Age and Illumination Mode (Technical Report Number FAA-AM-77-13). Oklahoma City, OK: Federal Aviation Administration Civil Aeromedical Institute.

Wulfeck, J.W., Weisz, A., and Raben, M.W. (1958). Vision in Military Aviation (WADC Technical Report 58-399). Wright-Patterson Air Force Base, OH: U.S. Air Force Wright Air Development Center.

Zeltner, K., Ratanaproeksa, P., Goldstein, S., Adams, S., and Green, P. (1988). Selected Abstracts and Reviews of the Legibility Literature (Technical Report UMTRI-88-22). Ann Arbor, MI: The University of Michigan Transportation Research Institute, June.

APPENDIX A
INDEX OF AUTHORS

APPENDIX B
INDEX OF TITLES