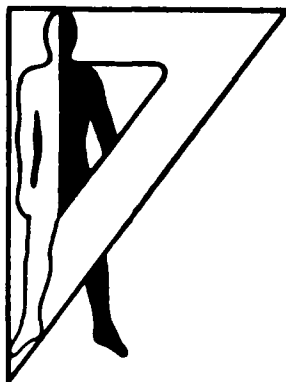


PHOCOPY RESOLUTION TEST CHART

AD-A180 272

DTIC FILE COPY

12



AD

Technical Memorandum 4-87

A LITERATURE REVIEW AND EXPERIMENTAL PLAN FOR RESEARCH ON THE
DISPLAY OF INFORMATION ON MATRIX-ADDRESSABLE DISPLAYS

Jennie J. Decker
Richard D. Pigion
Harry L. Snyder

Human Factors Laboratory
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

DTIC
ELECTE
MAY 18 1987
S D D

February 1987
AMCMS Code 612716.H700011

Approved for public release;
distribution unlimited.

U. S. ARMY HUMAN ENGINEERING LABORATORY
Aberdeen Proving Ground, Maryland

87 5 15 006

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Memorandum 4-87	2. GOVT ACCESSION NO. A180272	RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A LITERATURE REVIEW AND EXPERIMENTAL PLAN FOR RESEARCH ON THE DISPLAY OF INFORMATION ON MATRIX-ADDRESSABLE DISPLAYS	5. TYPE OF REPORT & PERIOD COVERED Interim Technical 16 Sept. 1985 - 16 Feb. 1987	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Jennie J. Decker, Richard G. Pigion, and Harry L. Snyder	8. CONTRACT OR GRANT NUMBER(s) DAAA15-85-R0066	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Human Factors Laboratory, Department of Industrial Engineering and Operations Research, Virginia Poly- technic Institute & State Univ., Blacksburg, VA 24061	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Human Engineering Laboratory Aberdeen Proving Ground, Maryland 21005-5001	12. REPORT DATE February 1987	13. NUMBER OF PAGES 139
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) U.S. Army Human Engineering Laboratory Aberdeen Proving Ground, Maryland 21005-5001	15. SECURITY CLASS. (of this report) Unclassified	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Unlimited Distribution		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Flat-panel displays, quality metrics, alphanumeric research, cartographic/symbolic research, literal image research, human factors engineering, visual displays		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) —This report summarizes the research on user performance and display quality metrics pertinent to flat-panel displays. The report is divided into five sections. Section 1 is a brief introduction. Section 2 provides a description of each of the current flat-panel technologies and compares the technologies based on nine display parameters. Section 3 summarizes research investigating the effects of various display variables on user performance. The research described in this section is divided into three		

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

▷ types: alphanumeric research, cartographic/symbolic research, and literal image research. Section 4 defines and discusses different quality metrics which have been used by researchers to predict user performance.

One purpose of this literature review was to determine variables requiring further investigation. Based on the findings, an experimental plan is presented in Section 5. An annotated bibliography of references pertinent to these topics is included at the end of the report. ←



Accession For	
NTIS CRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

A LITERATURE REVIEW AND EXPERIMENTAL PLAN FOR RESEARCH ON THE
DISPLAY OF INFORMATION ON MATRIX-ADDRESSABLE DISPLAYS

Jennie J. Decker
Richard D. Pigion
Harry L. Snyder

Human Factors Laboratory
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

February 1987

APPROVED:


JOHN D. WEISZ

Director

Human Engineering Laboratory

Approved for public release;
distribution unlimited.

Prepared for

U.S. ARMY HUMAN ENGINEERING LABORATORY
Aberdeen Proving Ground, Maryland 21005-5001

TABLE OF CONTENTS

	Page
SECTION 1 INTRODUCTION	9
Purpose	9
Objectives	9
SECTION 2 DISPLAY TECHNOLOGY	11
Display Descriptions	11
Cathode-Ray Tube (CRT)	11
Flat Panel CRT	11
Vacuum Fluorescent Display (VFD)	12
Plasma Displays	12
Electroluminescent Displays (EL)	12
AC Powder	12
AC Thin-Film	13
DC Powder	13
DC Thin-Film	13
Summary	13
Light Emitting Diodes (LED)	13
Liquid Crystal Displays (LCD)	14
Electrochromic Displays (ECD)	14
Electrophoretic Induced Displays (EPID)	15
Parameter Definitions	15
Physical Size and Configuration	15
Luminance	15
Spectral Emission	15
Element Size, Shape, Density	19
Contrast and Dynamic Range	19
Uniformity	19
Temporal Characteristics	21
Future Technology Projections	21
Comparisons of Display Technologies	22
SECTION 3 USER PERFORMANCE RESEARCH	31
Alphanumeric Legibility	31
Definition of Legibility	31
Dependent Measures	31
Type of Task	31
Display Luminance	33
Luminance Contrast	33
Ambient Illumination	34
Resolution	39
Continuous Displays	40
Discrete Element Displays	40
Matrix Size	41
Element Shape, Element Size, and Interelement Spacing	42
Element Shape	42
Element Size and Spacing	43

Contents Continued

	Page
Character Size	45
Stroke width	49
Symbol Width	50
Font	50
Dot versus Stroke Characters	51
Font Comparisons	51
Case	53
Viewing Angle	57
Character Orientation	57
Temporal Characteristics	58
Uniformity	59
Display Polarity	59
Cell and Line Failures	60
Chrominance	62
Color Contrast	62
Perceptual Problems	64
Cell and Line Failures on Multichromatic Displays	64
Areas in Need of Research	64
Relation to Display Technologies	65
Cartographic/Symbolic Research	67
Tasks and Dependent Measures	67
Resolution	67
Continuous Displays	67
Matrix Size	68
Cell and Line Failures	68
Chrominance	69
Areas in Need of Research	69
Relation to Display Technologies	69
Literal Image Research	70
Analog Imagery	70
Digital Imagery	71
Cell and Line Failures	75
Chrominance	75
Areas in Need of Research	75
Relation to Display Technologies	75
SECTION 4 IMAGE QUALITY METRICS	77
MTF-Based Metrics	78
Equivalent Passband (EP)	80
Equivalent Width (EW)	81
Squared Spatial Frequency (SSF)	81
Strehl Intensity Ratio (IR)	81
Perceptually Weighted System Metrics	81
Modulation Transfer Function Area (MTFA)	83
Gray Shade Frequency Product (GSFP)	84
Integrated Contrast Sensitivity (ICS)	84
Subjective Qualities Factor (SQF)	85
Perceived Modulation Quotient (PMQ)	85
Visual Capacity (VC)	85
Discriminable Difference Diagrams (DDD)	86
Just-Noticeable-Difference-Area (JNDA)	86
Displayed Signal-to-Noise Ratio	86
Hufnagel's Q ³	86
Signal-to-Noise (SN)	87

Contents Continued

	Page
Visual Efficiency (VE)	87
Information Content (IC)	87
Pixel Error Measures	88
Normalized Mean Square Error (MSE)	88
Point Squared Error (PSE)	88
Perceptual Mean Square (PMSE)	88
Image Fidelity (IF)	88
Structural Content (SC)	89
Correlational Quality (CQ)	89
Behavioral Validation of Image Quality Metrics	89
Monochrome Displays	90
Multichromatic Displays	97
Areas in Need of Research	98
SECTION 5 EXPERIMENTAL RESEARCH PLAN	101
Task Descriptions and Dependent Variables	101
Monochromatic Display Research	101
Monochromatic Alphanumeric Research	102
Experiment 1: Optimum Character Study	102
Experiment 2: Character Modification	102
Experiment 3: Failure Modes I	102
Experiment 4: Failure Modes II	103
Experiment 5: Element Characteristics	103
Experiment 6: Uniformity	103
Monochromatic Cartographic/Symbolic Research	103
Experiment 7: Optimum Character (Symbolic)	103
Experiment 8: Failure Mode Study (Symbolic)	103
Experiment 9: Element Characteristics (Symbolic)	104
Multichromatic Display Research	104
Multichromatic Alphanumeric Research	104
Experiment 10: Replication of Color	
Contrast Experiment	104
Experiment 11: Multichromatic Optimum	
Character Study	104
Experiment 12: Multichromatic Failure	
Mode Study	105
Experiment 13: Multichromatic Uniformity	105
Multichromatic Cartographic/Symbolic Research	105
Experiment 14: Multichromatic Optimum	
Character Symbolic	105
Summary	105
Quality Metrics Analysis	105
REFERENCES	107
ANNOTATED BIBLIOGRAPHY	117

List of Tables

Table	Page
1. Comparison of Physical Size Characteristics of the Display Technologies.	22
2. Comparison of Luminance Characteristics of the Display Technologies.	23
3. Comparison of Spectral Emission Characteristics of the Display Technologies.	24
4. Comparison of Element Size, Shape, Density Characteristics of the Display Technologies.	25
5. Comparison of Contrast and Dynamic Range Characteristics of the Display Technologies.	26
6. Comparison of Uniformity Characteristics of the Display Technologies.	27
7. Comparison of Temporal Characteristics of the Display Technologies.	28
8. Comparison of Display-Type Application Characteristics of the Display Technologies.	29
9. Comparison of Future Technology Projections of the Display Technologies.	30
10. Twelve Sources Reviewed by Banks, Gertman, and Peterson (1982).	32
11. Character Luminance Recommendations for Alphanumeric Legibility.	34
12. Modulation Recommendations for Alphanumeric Legibility.	37
13. Description of the Displays used by Duncan and Konz (1974).	38
14. Ambient Illumination Recommendations for Alphanumeric Legibility of Light Emitting Displays.	40
15. Matrix Size Recommendations for Alphanumeric Legibility.	42
16. Vertical Visual Angle Subtense (min of arc) (Snyder and Taylor, 1979).	48
17. Character Size Recommendations for Alphanumeric Displays.	49

List of Tables Continued

Table	Page
18. Comparison of Display Parameter Recommendations and Display Technologies.	66
19. Video System Linerates, Bandwidths, and Noise Levels Used by Snyder (1976).	71
20. Image Enhancement/Restoration Techniques Used by Chao (1983).	72
21. Usefulness of Restoration/Enhancement Processes for Improving Images Degraded by Blur and Noise (Chao, 1983).	74
22. References Which Relate Image Quality Measures to Observer Performance.	90
23. Correlations Between MTF-Based Image Quality Metrics and Observer Performance for Film Displays with Literal Imagery.	92
24. Correlations Between MTF-Based Image Quality Metrics and Observer Performance for CRT Displays with Literal Imagery.	93
25. Correlations Between Pixel Image Quality Metrics and Observer Performance for CRT Displays with Literal Imagery.	96
26. Correlations Between the MTF Image Quality Metric and Observer Performance for CRT Displays with Alphanumerics.	97
27. Pool of Predictor Variables, from Snyder and Maddox (1978).	98
28. Extended Predictive Equations, from Snyder and Maddox (1978).	99
29. Summary Matrix of Variables in Each Experiment.	106

LIST OF FIGURES

Figure	Page
Figure 1. Subjective colors within the chromaticity diagram.	17
Figure 2. Contour curves of maximum luminous reflectance for materials illuminated by CIE source C (daylight).	18
Figure 3. Seven-segment and starburst alphanumeric patterns.	19
Figure 4. An example of dot geometry for an alphanumeric character.	20
Figure 5. Percent of characters correctly identified as a function of size, blur, and contrast.	35
Figure 6. Speed of readout as a function of character size, blur, and contrast.	36
Figure 7. Mean values of recognition time for all digits at three levels of ambient illumination.	39
Figure 8. Effect of character/matrix size upon mean errors.	43
Figure 9. Element dimensions used by Burnett (1976).	45
Figure 10. Experimental design used by Burnett (1976).	46
Figure 11. Element sizes and shapes used by Maddox (1977).	47
Figure 12. Average distance in meters at which black and white numerals with different stroke-widths can be read.	51
Figure 13. Experimental levels used by Crook, Hanson, and Weisz, 1954, cited by Semple et al. (1977).	52
Figure 14. Maximum dot font in a 5 X 7 matrix used by Maddox et al. (1977).	54
Figure 15. Maximum angle font in a 5 X 7 matrix used by Maddox et al. (1977).	55
Figure 16. Lincoln/MITRE font in a 5 X 7 matrix used by Maddox et al. (1977).	56
Figure 17. Experimental design used by Abramson and Snyder (1984).	61
Figure 18. Combined blur and SNR (noise) used by Chao (1983).	73

List of Figures Continued

Figure	Page
Figure 19. Modulation transfer function.	80
Figure 20. Noise equivalent passband.	82
Figure 21. Visual contrast sensitivity function.	83
Figure 22. MTFA concept.	85

SECTION 1 INTRODUCTION

Purpose

The purpose of this report is to provide a comprehensive summary of the available human factors data concerning the effectiveness of the visual presentation of information on electronic displays. Past research has shown this effectiveness to depend upon several categories of variables, among them the symbolic representation of the information, the typography, and the information content. This review is particularly concerned with the variables concerning the presentation of information on matrix-addressable displays.

Such displays have constraints upon the placement and composition of characters and symbols which cannot be addressed by existing data on printed materials or cathode-ray tube (CRT) electronic displays. An example of these constraints is the fact that most matrix-addressed displays create characters from a series of lines or dots, rather than with continuous strokes as is the case with printed text, or some CRT displays. In addition, there is a fundamental difference between printed and electronic displays--namely, the occasional tendency for electronic (matrix-addressed) displays to fail locally. That is, some electronic displays will fail by having certain portions or elements of the display remain in the "on" or "off" state irrespective of the intended state of that display location. As the display failures increase in number, the display becomes logically less legible and therefore less usable. Unfortunately, data to support acceptability decisions and product quality assurance are generally unavailable, such that the user or purchaser is left with a decision to accept or reject a partially failed display with no supporting quantitative basis or data. Accordingly, this research effort is designed to remedy that problem.

Another major purpose of this report is to provide a comprehensive review of the various image quality metrics which have been designed to predict information transfer from electronic displays as a function of several objectively measured display parameters. Most of the quality metrics have been developed for spatially continuous displays. The data which exist for matrix-addressed displays (which are spatially discrete displays) are limited to predicting certain types of information transfer (e.g., from alphanumeric displays). Metrics which take the above mentioned display failures into account do not exist.

The literature review described above is used to form a basis for a comprehensive experimental plan to determine suitable design criteria for matrix-addressed displays.

Objectives

Specifically, the objectives of this research are as follows:

1. To provide additional needed quantitative experimental data on the effects of various types of matrix-addressable display failures on the ability of the user to obtain needed information from the display.
2. To provide quantitative data on the relationship between specific types of display presentation failures on information extraction for two types of monochromatic display content: alphanumeric and cartographic/symbolic.
3. To determine the quantitative effects of multicolor display content on the relationships indicated in (2) above.
4. To develop and recommend a quality metric that predicts information extraction performance as a function of the above variables that can be used by the U. S. Army for display evaluation, user performance prediction, and device quality assurance.

This report is organized into five sections. Section 2 contains brief descriptions of the various matrix-addressable (flat-panel) display technologies. Also included are comparisons of parameters common to all the displays.

Section 3 summarizes existing experimental data relating user performance to characteristics of flat-panel displays, failures of flat-panel displays, and design variables pertinent to flat-panel displays. Display parameters are discussed in terms of alphanumeric legibility, cartographic/symbolic research, and literal image research.

Section 4 reviews the existing and likely models of display quality pertinent to the design variables and failure modes of flat-panel displays. The appropriate formulae, original references, and limitations of the selected metrics are provided as well as are data for those metrics which have been behaviorally validated.

Section 5 describes the detailed experimental research plan which is designed to meet the objectives of this research program. The plan contains specifics of experimental designs, tasks, and independent and dependent variables.

An annotated bibliography is included at the end of the report.

SECTION 2

DISPLAY TECHNOLOGY

This section of the review summarizes the candidate flat-panel technologies. Although the conventional cathode-ray tube (CRT) and its several variants are neither solid-state nor flat-panel, the CRT is included in this discussion to serve as a baseline for comparison with the other technologies.

Each of the display technologies is described briefly. The purpose of these descriptions is to give some of the advantages and disadvantages of each display technology and to provide a simple description of its method of operation. Following these descriptions, 10 categories or parameters are defined which are useful in providing a comparison of the display technologies. These categories range from physical characteristics (size, configuration) through visual system pertinent variables (spectral emission, luminance, element size, element shape, contrast, uniformity, temporal characteristics). Also included are more subjective comments as to the utility of the technology for three categories of information presentation. In addition, a future technology projection is offered for each category.

Finally, comparisons of all the display technologies are made parameter by parameter. This information is presented in tabular form at the end of this section.

Display Descriptions

Cathode-Ray Tube (CRT)

The CRT dominates the market for a great majority of data and imaging applications. It is popular because it is relatively inexpensive compared to other display systems, has a long lasting familiarity with systems designers, and is extremely flexible. Advantages of CRTs are that they are available in a variety of sizes and shapes, provide gray scale and color, can have reasonably good resolution, can provide a storage capability, and can be addressed with both raster and stroke patterns. Some disadvantages are that the tube depth is equal to or greater than the display area, thereby giving it considerable bulk. Although it does have storage capability, it cannot store information at high luminance levels, and it has reduced detail contrast at high luminance.

Information on CRT capabilities is readily available from many sources (e.g., Sherr, 1979; Tannas, 1985). For comparison purposes the CRT will be listed in the summary tables at the end of the display descriptions.

Flat-Panel CRT

Although the conventional CRT has great flexibility as an information display, it has some substantial disadvantages. One of the major disadvantages is its depth. As the displayed image size is increased, so generally is the length of the tube. For this reason much effort has been placed on the development of the flat-panel CRT.

The concept of the flat-panel CRT can be illustrated by the Northrop Corporation's Digisplay™. The electron area source is a cathode which is less than 12 mm thick and consists of a number of cathode elements requiring fairly low power. The modulation plate controls the electron beam current from the cathode, much as the control grid does in a conventional CRT.

Many of the cathode techniques which have been developed for flat CRTs produce inadequate current output to achieve the desired luminance values. Because of this, several techniques have been developed (besides improving cathode output) to increase luminance. Among these are multiple beam addressing, electron multipliers, and storage techniques.

Beam positioning and modulation in flat-panel CRTs range from beam-deflection techniques, which are common to the traditional CRT, to matrix-addressed approaches. In the matrix-addressed versions, a control layer is used to selectively control the passage of electrons. The selected electrons then excite the cathodoluminescent phosphor screen.

In the matrix-addressed approach, the modulation plate is followed by a series of switching plates, each of which has an array of channels ("holes") which pass electrons. These switching plates accomplish two functions: (1) they keep the electron flow in well-defined channels or directions, and (2) they either pass or stop the flow of electrons in a given area by voltage addressing of each plate.

The flat CRT has advantages over other flat-panel approaches. Among these advantages are: (1) it uses a well-established technology derived from the conventional CRTs. (2) it uses

high-efficiency phosphors, (3) it can produce high luminance with good gray scale, and (4) the potential for achieving full color large size displays is good.

Vacuum Fluorescent Displays (VFD)

Vacuum fluorescent displays have been among the most successful of all flat CRT approaches. Contributing to its success has been the use of a patterned anode substrate combined with a low-voltage (50-100 volts) phosphor. Among the advantages of this technique are long life, pleasing appearance, rugged construction, and high luminance. The VFD is one of the lowest power and highest luminance light-emitting flat-panel displays currently available. Although originally developed to present alphanumeric, larger displays have recently been developed.

Plasma Displays

There are two types of plasma displays, one AC driven and the other DC driven. Both have been fabricated in alphanumeric readouts as well as in matrix-addressed panels for graphics and alphanumeric.

Plasma displays have one transparent (front) electrode, through which the display is viewed. The rear electrode can be black, reflective, or clear. When the rear electrode is clear, it is possible to rear-project an image on the display, thereby using the plasma display as overlay information on the projected image. This configuration is useful, for example, for map-type displays where a fixed high-resolution map is projected and the plasma display is used to indicate activity on the various regions of the map.

The basic mechanism of a plasma display is a gas filled volume across which an electrical field can be controlled. The electrical potential can cause the movement of an electron from one energy level to a lower energy level, simultaneously separating the electrons from the atoms. When a sufficiently large number of atoms have lost at least one electron, the gas is said to be in its ionized state. This ionization process produces the cathode glow resulting in light emission.

In the DC-driven configuration, the electrodes are located inside glass plates, in direct contact with the gas-filled center cavities. The AC-driven display has the electrodes separated from the gas. Both types of plasma displays are matrix-addressable.

Recently, plasma displays have been developed that use both the AC and DC methods. The purpose of these displays is to combine the best features of both techniques into higher performance displays (Weber, 1985).

Electroluminescent Displays (EL)

Electroluminescence (EL) is the emission of light from a phosphor after application of an electric field. EL displays are made up of either phosphor powders or thin-film layers of polycrystalline materials. They may be excited by either AC or DC current, thus providing four generic display types: AC powder, AC thin-film, DC powder, and DC thin-film. The phosphor material most commonly used in both powder and thin-film displays is zinc sulfide (ZnS) activated by copper (Cu), although other phosphors and activators are also used (Lehmann, 1980; Snyder, 1980; Tannas, 1985).

The basic construction of an EL display or panel places the phosphor between a pair of row electrodes and a pair of transparent column electrodes. The transparent electrodes are placed against the glass substrate. For powder EL panels the phosphor powder is often actually sprayed or screened onto the glass substrate. Also, for DC EL panels the phosphor cannot be continuous from row to row or shorting will occur because the DC excited phosphor is conductive (Tannas, 1985).

With the row-column electrode configuration, light is only emitted where the two electrodes overlap (element). Above the threshold required to excite the element, an increase in voltage causes the phosphor to glow proportionally brighter, producing grayscale (Graff, 1985). The advantages and disadvantages of the four EL configurations will be briefly discussed.

AC Powder. AC powder EL displays are used for applications which require continuous low-luminance such as transillumination of panels, keyboards, or other displays (Tannas, 1985). Long life is possible at low luminance levels (7 ft-l) but not at the moderate or high luminances that are required for alphanumeric displays. When driven at high luminance there is an exponential decay resulting in a 50% reduction in display luminance after only 1,000 hours of operating life (Tannas, 1985).

Howard (1981) and Tannas (1985) have pointed out that it is difficult to construct complex matrix-addressed displays with AC powder due to their low discrimination ratio (lack of nonlinearity, the more nonlinear the display in luminance response the more compatible it is to

matrix addressing; Snyder, 1980). Improvement in nonlinearity can be made with the use of thin-film transistors added at each pixel.

Another disadvantage is that contrast ratios are low in moderate to high ambient illumination because the powder reflects ambient light. An absorbing filter on the front of the display helps to reduce this problem; however, higher display luminances are then required and viewing angles are also reduced.

AC Thin-film. According to Tannas (1985), EL has become a viable technology due to the use of thin-film phosphors. AC thin-film is currently the most promising EL display type and has been used for several commercially available systems. AC thin-film EL displays have been shown to have long life, to be sunlight readable, have high luminance, and better discrimination ratios than the other EL technologies (depending on the phosphor used).

Display luminance is controlled by varying the refresh rate and pulse width. For each refresh frame a pixel gives two pulses of light, a characteristic which allows greater flexibility for controlling flicker and luminance (Tannas, 1985). Also, AC thin-films are highly nonlinear which, as previously pointed out, is necessary for matrix addressability. Sharp reported the life of a thin-film phosphor to be over 20,000 hours with no aging effects, although "large panels" typically have a 30% luminance reduction after 10,000 hours (Tannas, 1985). This is competitive with CRT phosphor life.

The largest disadvantage of this type of EL display is that as the number of lines to be refreshed is increased, the pulses must be shortened to avoid flicker, thus leading to high voltages and high cost driving systems (Howard, 1981). As driving voltage increases, the life of the display decreases (Snyder, 1980).

DC Powder. DC powder displays are more easily matrix-addressed than AC powder displays due to the higher discrimination ratio (Tannas, 1985). The contrast ratio is also better for DC powder than AC powder because the luminance of DC powder is proportional to the sixth power, while luminance is proportional to the third power for AC powder (Snyder, 1980; Tannas, 1985). Applications for this technology include automotive panels, and 80- or 256- character displays (Tannas, 1985).

The disadvantages of this technology include limited resolution compared to thin-film and AC powder due to the thickness of the powder used; poor contrast in high ambients due to reflection off the powder; and continuously increased voltage to maintain luminance until electrical breakdowns destroy the phosphor film (Tannas, 1985).

DC Thin-film. Although DC thin-film is one of the oldest configurations, it is far behind in development. An advantage of this configuration is low operating voltages; however, according to Tannas (1985), the need for low voltage disappeared with availability of high voltage drivers. The problem with this configuration is the tendency towards catastrophic failures (Howard, 1981; Tannas, 1985).

Summary. In general, EL displays require high voltages, as much as 100 times that of liquid crystals of the same size. This requirement rules out the possibility of battery operation (Graff, 1985). EL displays have higher contrast and better resolution than liquid crystal displays (LCDs), allow for wider viewing angles, and they are far less bulky than CRTs.

ELs are currently all monochromatic. A wide range of colors is available depending upon the phosphor used. Planar Systems is currently working on an Army contract to develop multichromer displays, and has developed several experimental prototypes. One approach is to use three phosphor layers (red, green, and blue) with a separate matrix of electrodes for each layer. The intensity of the primaries at each element would determine the hue (Graff, 1985). Problems encountered are that the driver electronics are considerable and different color phosphors do not glow with the same efficiencies. Addressing red and blue phosphors more often than green is one possible solution but even more drive electronics are then necessary.

Light Emitting Diodes (LED)

The light emitting diode (LED) is a form of electroluminescence. Light is emitted from these devices after application of an electric field. An LED is a single semiconductor device consisting of a single p-n junction. The device emits light after voltage application to the forward-biased p-n junction (Craford, 1985).

LEDs are commonly used for applications such as calculators, watches, and instrument panels due to their high reliability (LEDs do not have a tendency towards catastrophic failures), high luminance, low power, low cost, and compatibility with integrated circuit technology (Snyder, 1980). Unfortunately, when single LEDs are used to make up an x-y array for large screen applications the power requirements become exorbitant. Also, luminance of an LED increases linearly with increases in current; therefore, in high ambients when higher luminances are required for contrast, power

requirements become unacceptable, especially as the number of elements in a matrix array increases (Snyder, 1980).

LEDs have very sharp rise and decay times, on the order of 10-10,000 ns range (Goodman, 1974). Thus, high refresh rates are required so that the display does not flicker. Refresh rates range from 400-1,000 Hz for these devices (Snyder, 1980).

The colors available for LEDs are currently limited to red, green, orange, and yellow. Several colors may be used on one display.

Craford (1985) stated that there are no large screen LED displays available on the market, although they have been prototyped. Compared to other technologies large array displays are still uneconomical and impractical.

Liquid Crystal Displays (LCD)

Liquid crystal displays (LCDs) do not emit light after application of a voltage. Instead they control or modify ambient illumination by scattering light, or modulating optical density, or changing color (Goodman, 1974). They have been termed "passive" displays because they do not emit or generate light.

LCDs are a popular technology and a great deal of research is being conducted trying to optimize LCDs. LCDs are constructed by placing the liquid crystal material between two glass plates which are partially covered with conductive coatings. One side must be transparent.

There are several categories of LCDs, defined by the molecular organization and operating characteristics. These characteristics will not be discussed here. Readers are referred to Penz (1985) and Goodman (1974) for a discussion of these properties. This section will focus on general advantages and disadvantages of the technology.

The major advantage of the LCD is that it requires very little power to operate, thereby allowing for battery operation, which is a necessity for portability. The most common applications for LCDs have been calculators, watches, and other portable applications. Another advantage is that since LCDs modulate ambient light, they are readable under high ambient conditions including sunlight.

According to Penz (1985) a major disadvantage is that they are limited to almost nonexistent matrix addressing capability. This capability is required for application of large screen displays and high information content displays (such as literal images). One difficulty in matrix addressing is due to the long rise and decay times of LCDs (Goodman, 1974; Snyder, 1980). The rise and decay times are dependent upon the fluid's viscosity and are affected by temperature, becoming longer at higher temperatures. Different types of LCDs have different rise and decay times.

Matrix addressed LCDs have been constructed, although they generally have poor contrast. As display size increases contrast deteriorates (Aldersey-Williams, 1985). Data General introduced the first multiplexed LCD personal computer in 1984, the Data General/One™. It is a 640 X 640 element display with battery operation for portability. The contrast was only 3:1 (Aldersey-Williams, 1985). Similar displays have been used for other portable computers.

Active matrix addressing is being used for large screen LCDs. A semiconductor is placed at all row and column intersections so that the voltage signal only affects the intersected element. Active matrix addressing has allowed multicolor and gray scale displays (Aldersey-Williams, 1985; Laycock, 1985a). A 480 X 480, 100-mm diagonal color television was designed by SUWA Seikosha (*Information Display*, 1985). Panelvision markets a 192 X 128 panel, which is priced 10 times as high as a CRT (Aldersey-Williams, 1985). Although matrix addressing is difficult, it has been accomplished and it appears that technological advances will continue.

Another problem with LCDs is off-axis viewing limitations. For twisted nematic LCDs, the contrast varies with the angle of view relative to normal (0 degrees) and relative to the angle of incidence of the ambient illumination. With higher driving voltages, greater contrast is obtained farther off-axis (Snyder, 1980). Different types of LCDs have different viewing angle limitations. It should be noted that multiplexing further reduces the viewing angle (Sutton & Powers, 1984).

Electrochromic Displays (ECD)

Electrochromic displays (ECDs) are nonemissive light modulating devices like LCDs. An ECD is similar to a battery with one electrode serving as the display (Penz, 1985). The transparent electrode absorbs a selected portion of the visible spectrum upon application of an electric field. The color of the "on" portion is dependent upon the material used to fabricate the display.

Like an LCD, the advantages of the ECD are low voltage and sunlight readability. An ECD has better contrast than an LCD and contrast does not depend upon the angle of view (Penz, 1985).

ECDs have inherent memory. When turned on they remain on for days after the voltage is removed or until they have been bleached by application of a reverse voltage. Bleaching takes about 1 second while rise time is on the order of seconds (Snyder, 1980). Due to slow response times matrix addressing is difficult although it has been accomplished (Nicholson, 1984; Penz, 1985). Matrix addressing increases the power consumption.

Applications of this technology are still limited. Alphanumeric readouts such as watches and calculators are available. Due to their slow response times, watches can display minutes and hours but not seconds (Penz, 1985). For that reason, ECDs are not as common as LCDs for these applications. ECDs are still unsuitable for graphic or literal image displays, and the research in this area is slow. There are not many companies interested in advancing this technology.

Electrophoretic Induced Displays (EPID)

The EPID is a nonemissive (light modulating), rather than an emissive (light emitting) display. It results from the process of electrophoresis, which is the movement of charged particles suspended in a liquid by the application of an electric field. The pigmented particles are selected to be a different color or optical density than the suspending liquid, so that the migration to the front surface of the display cell permits the observer to "see" the particles, whereas migration to the rear surface of the display causes the observer to see only the suspending liquid. Selection of colors or optical densities of the pigmented particles versus the suspending liquid determines the contrast or chromaticity of the EPID.

Like many other solid-state displays, the EPID is essentially a transparent sandwich, with the front and rear plates coated with conducting electrodes. The cavity created by spacers between the two transparent electrodes is filled with a fluid composed of a small pigmented particle suspension in a dense liquid.

The application of an electric field across the electrodes causes the particles to migrate toward one or the other electrode. The rate of migration of the particles depends on several factors, among them the particle size, the cell thickness, and the field voltage.

Parameter Definitions

Physical Size and Configuration

This category describes the typical size and the range of physical sizes over which the display type can or may be fabricated. In some cases, the discussion refers to commercially available sizes, in other cases to potentially available sizes. In a couple of cases, limits to size are noted, as constrained by the inherent technology characteristics.

In addition, the basic physical configuration(s) of each technology is described so that the design limitations of each device may be evaluated parameter by parameter. No effort is made to present detailed quantitative design trade-offs. The present discussion is intended to reveal the available design data which may be of importance in the selection of the experimental variables of interest in the present research program.

Luminance

The visual system is not equally sensitive to all wavelengths of visible radiant energy; therefore, the radiant energy must be weighted by the sensitivity of the eye to that wavelength. This sensitivity weighting function is termed the photopic luminosity function. The eye is most sensitive in the middle or green section of the visible spectrum, and least sensitive at the extreme red (long-wave) and blue (short-wave) ends.

The weighting of radiant energy by the photopic luminosity function yields the physical measure of luminance expressed in candelas/square meter (cd/m^2). Other units commonly used are foot-Lambert (ft-L), millilambert (mL), and others. The cd/m^2 is commonly referred to as the nit. One foot-Lambert equals $3.426 \text{ cd}/\text{m}^2$.

Brightness is a subjective perception and not a physical measure or property of a display surface and cannot be measured in physical units. "Brightness" is affected by spectral emission of the display and the surround, the visual adaptation state of the observer, and the luminance of both the display and surround.

Spectral Emission

The human visual system is not equally sensitive to all wavelengths of visible light energy. Accordingly, wherever possible the spectral emission is given in either radiant or luminous energy

per unit wavelength. In keeping with current scientific usage, wavelength is expressed in nanometers (nm).

The color spectrum may be described in several ways. For example, visible light energy can be described in electromagnetic energy space as that portion of the electromagnetic wavelength (or frequency) domain to which the eye is sensitive, ranging approximately from 380 to 720 nm. Very narrow wavelength bands produce "pure" colors. Any visually dominant wavelength can be synthesized from other colors, in accordance with the CIE Standard Observer chromaticity diagram (Snyder, 1980).

The CIE Standard Observer allows a chromatic stimulus to be specified in a standardized form. To define a chromatic stimulus, the tristimulus values X , Y , and Z are computed from the spectral radiance of the stimulus. These tristimulus values of X , Y , and Z are the amounts of the red, green, and blue primaries, respectively, which would be required to match the stimulus color. From these tristimulus values, chromaticity coordinates are computed. These coordinates are defined as $x = X/(X + Y + Z)$, $y = Y/(X + Y + Z)$, and $z = Z/(X + Y + Z)$. For convenience, the x, y, z chromaticity coefficients which define all spectral colors are conventionally plotted in x, y coordinates, noting that $x + y + z = 1$. Subjective colors existing in various parts of the CIE space are labeled in Figure 1.

It is often convenient to think of luminance as a dimension orthogonal to the x, y chromaticity diagram. For emissive displays, luminance can be independent of the x, y coordinates of the display, subject only to the emissive properties of the display device. For a reflective display (e.g., liquid crystal), color is obtained by selective absorption or transmission. Thus, the maximum luminance (or maximum reflectance) usually occurs with white light, assuming a white light ($x = y = z$) ambient source. For selectively absorbing displays, greater absorption produces "purer" colors, at the expense of reduced luminance or reflectance. The maximum possible reflectance, as a function of x, y coordinates, is shown in Figure 2.

These relationships will be referred to in later sections.

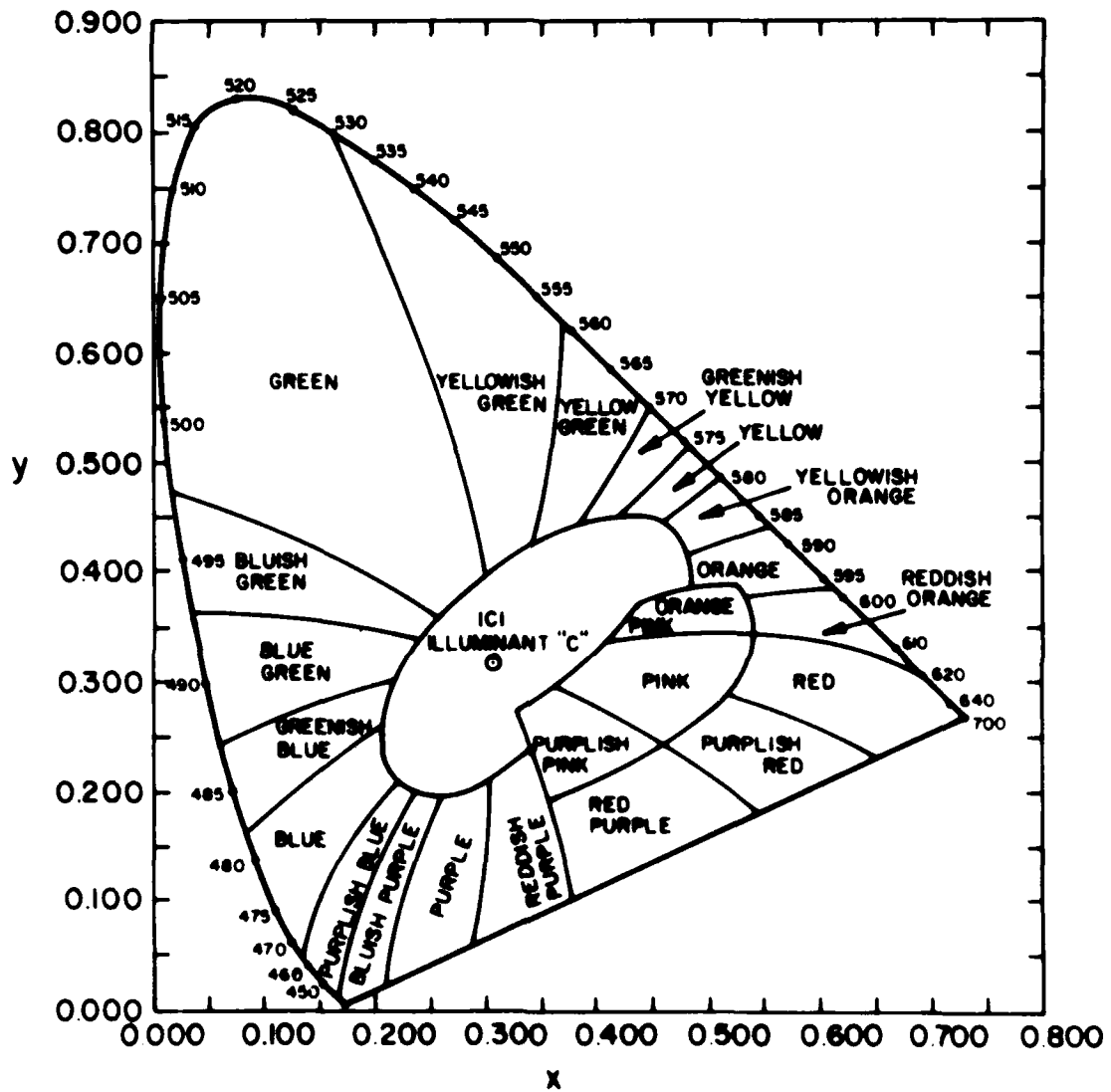


Figure 1. Subjective colors within the chromaticity diagram.

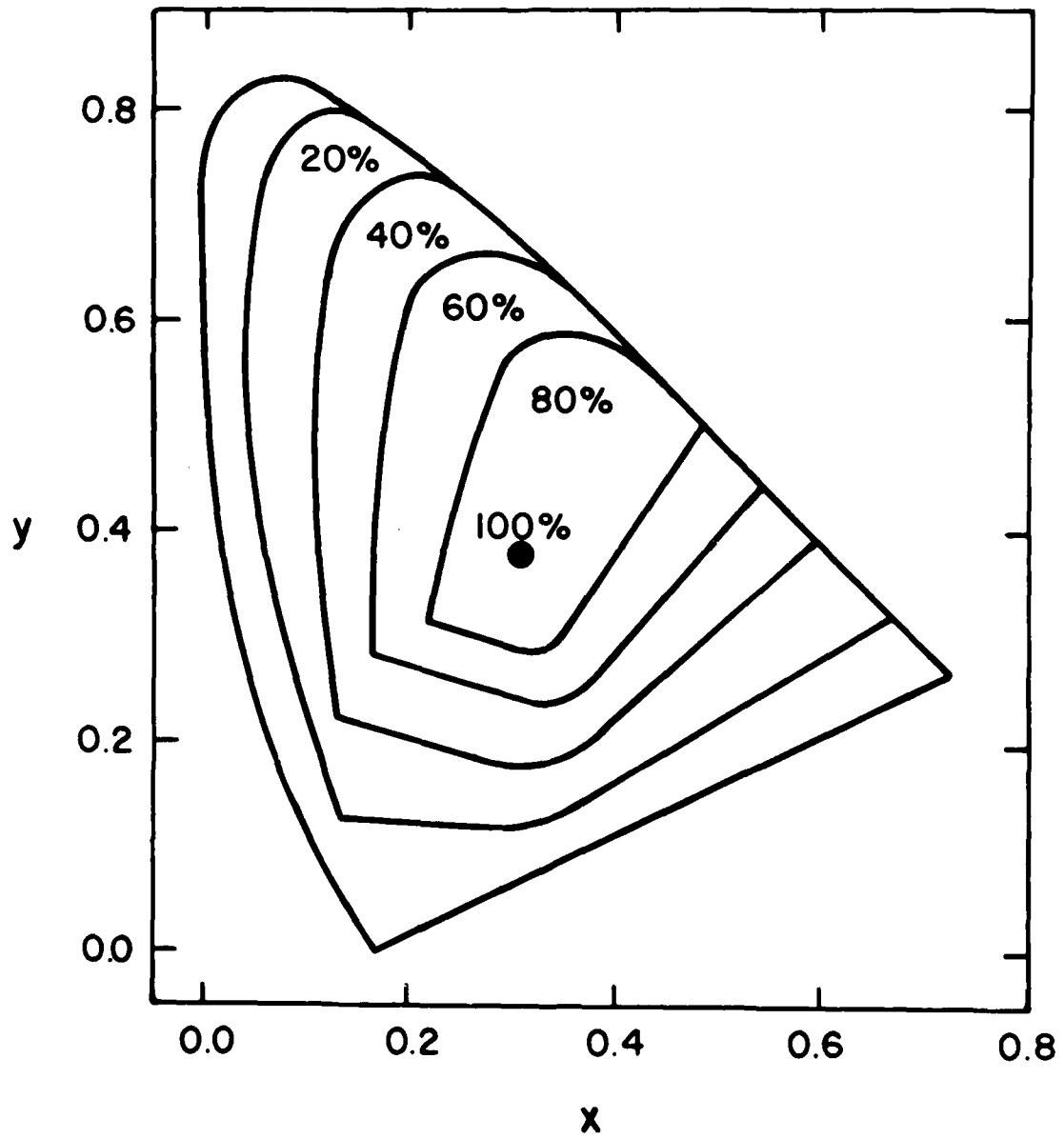


Figure 2. Contour curves of maximum luminous reflectance for materials illuminated by CIE source C (daylight).

Element Size, Shape, Density

Flat-panel displays are generally segmented in one of two forms. Alphanumeric readouts are often designed from fixed line segments such as a seven-segment or starburst pattern (Figure 3), in which each segment is addressed separately. The other form is that of an element matrix. In this form elements are arranged in an X-Y array. Selection of individual elements in the array allows creation of alphanumerics, symbols, lines, solid shaded areas, or pictorial information such as on a television screen. Readability and legibility of X-Y matrix-addressed displays are affected by the element size, shape, and spacing between elements. Element size is typically expressed in diameter, or length and width. Element shapes are specified by appropriate terms such as circular, square, Gaussian, etc. Spacing between elements is specified either as edge to edge or center to center. Figure 4 is an example of dot geometry for an alphanumeric character.

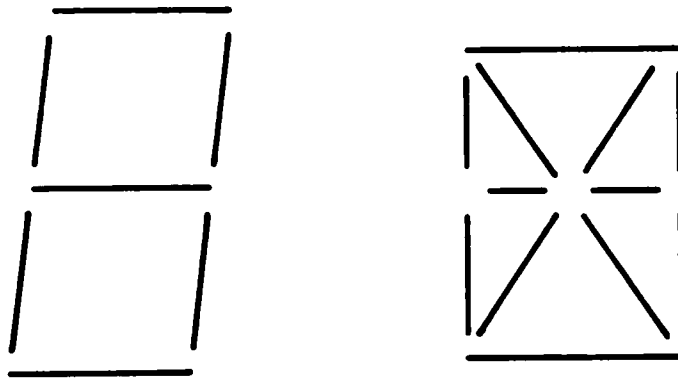


Figure 3. Seven-segment and starburst alphanumeric patterns.

Contrast and Dynamic Range

While display element luminance is important in display design, an equally important parameter is the contrast between any "on" element and its "off" background. Unfortunately, the literature contains many definitions of "contrast." If the maximum or "on" luminance is symbolized as L_{\max} and the background or "off" luminance is indicated by L_{\min} , then the following relationships hold:

$$\text{Modulation (M)} = (L_{\max} - L_{\min}) / (L_{\max} + L_{\min}), \quad (1)$$

$$\text{Contrast Ratio} = (L_{\max} / L_{\min}) = (M + 1) / (1 - M), \quad (2)$$

$$\text{Dynamic Range} = L_{\max} - L_{\min} = L_{\max} (2M) / (M + 1), \text{ and} \quad (3)$$

$$\text{Relative Contrast} = (L_{\max} - L_{\min}) / L_{\min} = (2M) / (1 - M). \quad (4)$$

In general modulation and contrast ratio are the most useful and most used terms.

Uniformity

Uniformity is best defined by its absence, or by nonuniformity. Goede (1978, cited by Snyder, 1980) defined three types of nonuniformity. Large area nonuniformity refers to the gradual change in luminance from one area of the display to another, for example, the change in luminance from the center of the display to the edge. Large area nonuniformity exists on most displays (Snyder, 1985). Small area nonuniformity refers to luminance (or color) changes from element to element. Edge discontinuity refers to changes in luminance or color over an extended boundary. While this

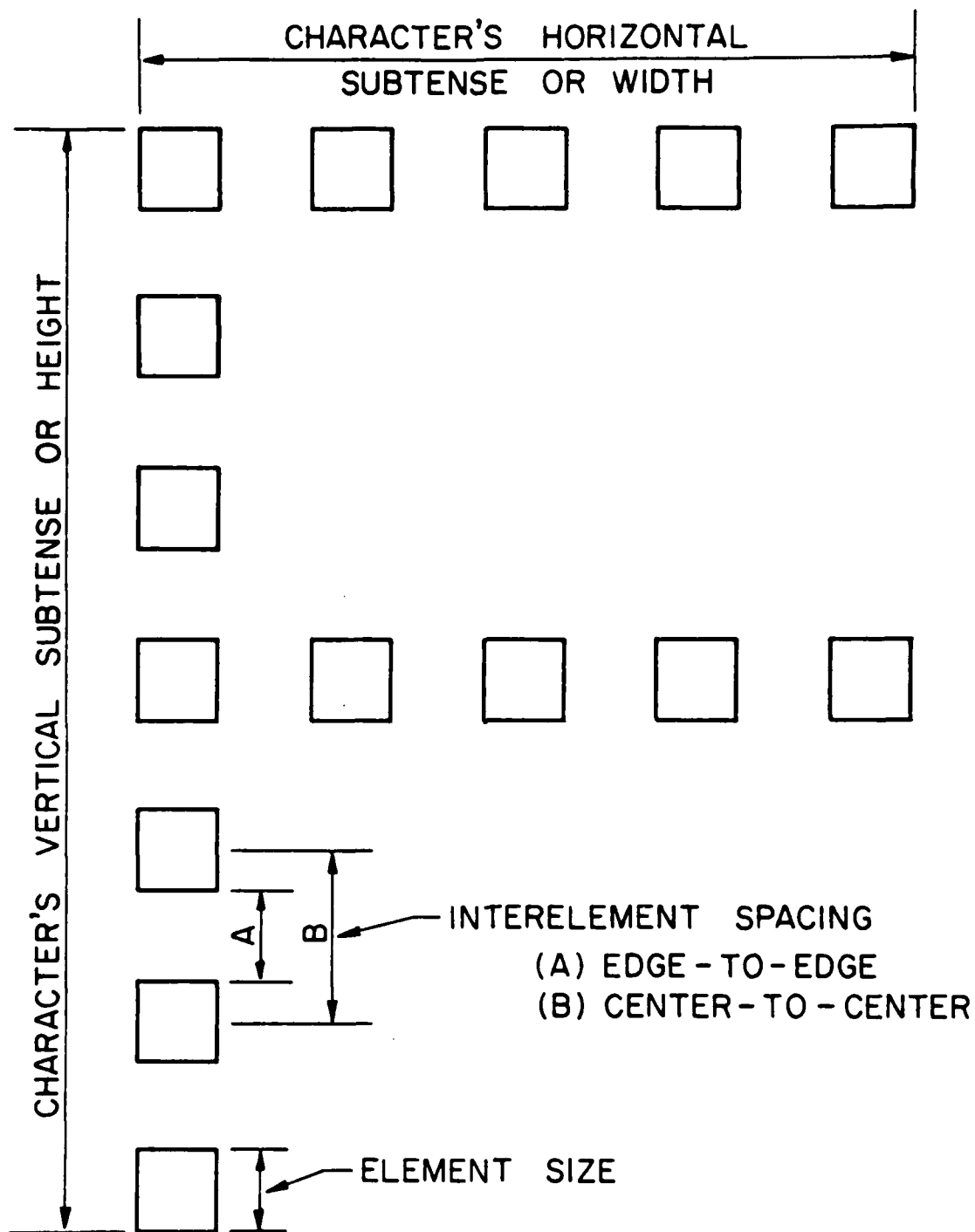


Figure 4. An example of dot geometry for an alphanumeric character.

classification of nonuniformity is helpful, still to be defined in the scientific literature are the terms large area, small area, and changes in luminance.

Temporal Characteristics

Some flat-panel displays have inherent memory so that when an element is turned on it remains on until turned off. Most technologies, however, have display elements which require periodic refreshing to avoid the perception of flicker. To determine the required refresh rate to avoid flicker, the rise and decay time of the luminance of the device must be known.

Rise time refers to the time period required by the device to reach maximum luminance after the application of a squarewave "on" pulse or command. It is typically measured in microseconds ($1 \mu = 10^{-6}$ s.) or milliseconds ($1 \text{ ms} = 10^{-3}$ s.)

Decay time is the time, following cessation of the "on" pulse or command, for the luminance to reach 10% of its maximum value. It is also measured in microseconds or milliseconds.

Future Technology Projections

Where possible, information is given on the future directions of research and development. Areas of improvement critical to meeting various application requirements and performance criteria are noted.

Comparisons of Display Technologies

As mentioned in the introduction to Section 2, the Tables 1-9 provide a comparison of the displays parameter by parameter.

Table 1
Comparison of Physical Size Characteristics
of the Display Technologies

Display Type	Display Size (Typical)	Display Depth (Typical)
CRT	≤91 cm diag.	1 to 4 x display diag.
Flat-Panel CRT	75 x 100 cm	10 cm
Vacuum Fluorescent	(10.2 cm) ²	1.99 cm
Plasma Discharge	140 x 140 cm (max.)	12 cm
EL	9.6 x 19.2 cm	1.905 cm
LED	12 x 16 cm	1 cm
LCD	12.15 x 24.3 cm	.012 cm
EC	unknown	.1-.2 cm
EPID	15 x 30 cm	.1-.2 cm

Table 2
 Comparison of Luminance Characteristics of the
 Display Technologies

Display Type	Maximum Luminance (cd/m ²)	Minimum Luminance (cd/m ²)	Dependent on Resolution
CRT	34,000 300 typical	1-2	yes
Flat-Panel CRT	820	1-2	yes
Vacuum Fluorescent	750 270 typical	1-2	yes
Plasma Discharge	600	0	no
EL	3,400	0	no
LED	34,000	0	no
LCD	n/a	n/a	no
EC	n/a	n/a	no
EPID	n/a	n/a	no

Table 3
Comparison of Spectral Emission Characteristics of the
Display Technologies

Display Type	Dominant Wavelength	Spectral Dispersion	No. of Discriminable Colors Available ^a
CRT	varies with phosphor	varies with phosphor	≤ 20 with 3-gun CRT
Flat-Panel CRT	varies with phosphor	varies with phosphor	≤ 20 with triad dots
Vacuum Fluorescent	varies with phosphor	varies with phosphor	≤ 20
Plasma Discharge	585 nm (neon) others less	varies with phosphor/gas	≤ 20 with full color 1 otherwise
EL	585, 525 nm varies with phosphor	100 nm	approximately 7
LED	650, 632, 590, 560, 490 nm	wide, continuous	5
LCD	varied	unknown	unknown
EC	varied	varied	unknown
EPID	varied	unknown	unknown

^a Assumes absolute categorization under typical ambient illumination.

Table 4
 Comparison of Element Size, Shape, and Density Characteristics
 of the Display Technologies

Display Type	Element Size Minimum, mm	Element Shapes	Element Density
CRT	0.07 at 2.35σ	Gaussian	variable
Flat-Panel CRT	0.35 (est.)	Gaussian	to 3.15/mm
Vacuum Fluorescent	0.125	Gaussian	?
Plasma Discharge	0.25	variable	to 3.27/mm
EL	$(.279)^2$	selectable	3.6/mm
LED	.300 x .250	round, square	4/mm
LCD	.180 x .135	selectable	20/mm
EC	$(3.175)^2$	selectable	.315/mm
EPID	?	?	?

Table 5

**Comparison of Contrast and Dynamic Range Characteristics of
the Display Technologies**

Display Type	Maximum Modulation	Dependent on Ambient Illumination	Light Emitter or Light Modulator
CRT	98%, at low luminance and low ambient	yes	emitter
Flat-Panel CRT	98%	yes	emitter
Vacuum Fluorescent	98%	yes	emitter
Plasma Discharge	95%	somewhat	emitter
EL	92%	somewhat	emitter
LED	96%	somewhat	emitter
LCD	96%	yes	modulator
EC	90%	yes	modulator
EPID	94%	yes	modulator

Table 6
Comparison of Uniformity Characteristics of the Display Technologies

Display Type	Small Area	Large Area	Image Geometric Stability
CRT	good	fair, 50% rolloff	fair
Flat-Panel CRT	fair	fair to good	good
Vacuum Fluorescent	good	good	good
Plasma Discharge	good	good	very good
EL	fair	fair	very good
LED	good	poor	very good
LCD	good	fair	very good
EC	probably good	unknown	very good
EPID	probably good	unknown	very good

Table 7

Comparison of Temporal Characteristics of
the Display Technologies

Display Type	Rise Time	Fall Time	Inherent Memory	Refresh Requirements
CRT	1 μ s to 1 ms, depends on phosphor	1 μ s to > 100 s, depends on phosphor	typically not, except for storage CRTs	varies with phosphor
Flat-Panel CRT	same as CRT	same as CRT	same as CRT	same as CRT
Vacuum Fluorescent	same as CRT	same as CRT	same as CRT	same as CRT
Plasma Discharge	100 ns	2 μ s	yes	50-60 Hz
EL	1 ms	0.1 ms to 1.5 ms	yes	60 Hz
LED	10-1000 ns	10-1000 ns	no	400-1000 Hz
LCD	50-300 ms	100-400 ms	yes	none
EC	0.1-1.0 s	0.1-1.0 s	yes	none
EPID	10-100 ms	10-100 ms	yes	none

Table 8

Comparison of Display-Type Application
Characteristics of the Display Technologies

Display Type	Single Alphanumeric	Matrix (graphic)	Matrix (TV)	Gray Scale
CRT	possible, but not practical	yes	yes	yes
Flat-Panel CRT	yes	yes	yes	yes
Vacuum Fluorescent	yes	yes	yes	yes
Plasma Discharge	yes	yes	yes	yes
EL	yes	yes	monochrome only	yes
LED	yes	available, but too costly	prototyped, but too costly	yes
LCD	yes	yes	yes	yes
EC	yes	no	no	no
EPID	yes	yes	doubtful	yes

Table 9
 Comparison of Future Technology Projections
 of the Display Technologies

Display Type	Mature Technology	Major Improvements Required for Wide-spread Usage	R&D Trends
CRT	yes	none	better uniformity, resolution
Flat-Panel CRT	moderately	color	full color
Vacuum Fluorescent	yes	size	full color graphics
Plasma Discharge	yes	color	color resolution
EL	yes (monochrome)	color, luminous efficiency	color, luminous efficiency
LED	yes	uniformity, cost	color, luminous efficiency
LCD	yes	rise/fall times, angular viewing	response times, addressing
EC	no	response times, threshold	response times, threshold
EPID	no	response times, addressing	response times

SECTION 3

USER PERFORMANCE RESEARCH

The purpose of this section is to review literature relating user performance to the various characteristics of flat-panel displays. The largest proportion of the user performance literature focuses on alphanumeric legibility. Based on this literature, several guidelines pertaining to alphanumeric legibility and solid-state displays have evolved. The proportion of research investigating user performance with cartographic/symbolic or literal image displays is quite small in comparison to the alphanumeric research.

This section of the literature review has been divided into three subsections: alphanumeric legibility research, cartographic/symbolic research, and literal image research. Display parameters and their interactions will be reviewed for their effects upon user performance. Findings will be related to current display technologies.

Alphanumeric Legibility

A considerable amount of research has been conducted on alphanumeric legibility, and several recommendations for designing alphanumeric displays have evolved. This section discusses the literature in this area by display parameter. When recommendations have been offered they are presented.

Banks, Gertman, and Peterson (1982) compiled various parameter recommendations from 12 international sources. Many of these recommendations are presented in this section. Table 10 identifies the 12 sources. Throughout this report sources will be referred to by the acronyms listed in Table 10. Recommendations from other sources will also be presented in this review.

Definition of Legibility

Cornog and Rose (1967) pointed out that several terms are used in legibility research, including legibility, readability, perceptibility, and visibility. They state that legibility includes all of these terms and define legibility as referring "to the characteristics of printed, written, or other displayed meaningful symbolic material which determine the speed and accuracy with which the material may be read or identified."

Dependent Measures

Actually the definition of legibility really depends upon the measures used in the research. The most common dependent measures are response time and accuracy (number of errors or correct identifications). Tachistoscopic recognition and threshold visibility have also been described as dependent measures (Semple, Heapy, Conway, & Burnett, 1971; Snyder & Taylor, 1979). Both these measures, however, draw upon the use of response time and/or accuracy data. Subjective questionnaires have also been employed and even visually evoked responses (VERs) have been used (O'Donnell & Gomer, 1976).

Snyder and Taylor (1979) evaluated the sensitivity of four response measures commonly used in alphanumeric legibility research. Character size, luminance, and viewing distance were the display parameters manipulated. Recognition accuracy, response time, tachistoscopic recognition accuracy, and threshold visibility were the response measures investigated. (For tachistoscopic recognition exposure time instead of viewing distance was used.) Findings indicated that recognition accuracy was the most sensitive response measure. It was felt that response time provided important information, while tachistoscopic recognition was insensitive. The insensitivity of this measure may have been due to the short viewing distance and long exposure times. Threshold visibility was not directly comparable to the other measures because it was determined using the accuracy data; however, the data with this measure were found to agree with the recognition data. It appears that researchers are generally using the most sensitive measures.

Type of Task

Performance results yielded by response time or accuracy measures may be different depending upon the type of task employed. Tasks commonly used are letter recognition, word recognition, and

Table 10

Twelve Sources Reviewed by Banks, Gertman, and Peterson (1982)

TUB	Technical University, Berlin.
DCIEM	Defense and Civil Institute of Environmental Medicine
DIN	Draft DIN Standard. DIN is a German standard organization.
SNBOSH	Swedish National Board of Occupational Safety and Health.
VDT	Cakir, A., Hart, D. J., and Stewart, T. F. M. The VDT Manual, Inca-Fiej Research Association, Darmstadt, F. R. G., 1979.
GREV	Groupe de Recherche sur les Ecrans de Visualization.
U of L	University of London
IBM	International Business Machines Incorporated.
EG&G	EG&G Idaho, Inc., Idaho National Engineering Laboratory.
NIOSH	National Institute of Occupational Safety and Health
MIL-STD-1472B	U. S. Military Standard 1472B, 1981.
BRN	British Royal Navy

reading performance, for example by Tinker's Reading Test. Random and structured search tasks have also been used (Burnette, 1976) as has Sternberg's Test (O'Donnell & Gomer, 1976; Peters & Barbato, 1976). It is logical to assume that performance using a letter recognition task will be different than performance using a word recognition task, especially if the redundancy of the English language is considered (Albert, 1975). This is an important consideration when studying the effects of dot or line failure (or degradation in general) because, while subjects may not be able to recognize a single character, enough information may be available to recognize a word.

Albert (1975) evaluated contextual and noncontextual characters on performance. The display parameters of character sizes (2.64, 3.05, 4.79, and 5.44 mm) and display luminance (8, 24, and 66 cd/m²) were also investigated. Anagrams (scrambled words) and the unscrambled words were presented tachistoscopically. The dependent measure was the number of correctly recalled letters in their correct locations. Mean word score minus mean anagram score was used to evaluate the advantage of contextual over noncontextual stimuli. A significant interaction between character size and luminance was found. A significant difference existed between the highest luminance level and the two lower levels at the smallest character size (2.64 mm). The effect of character size was significant at all luminance levels but differed depending upon the luminance level. It was concluded that presentations of contextual letters will improve performance over noncontextual letter presentations under degraded conditions of low luminance and small character sizes.

These findings point out that the term readability should probably be considered separately from the definition of legibility. McCormick and Sanders (1982) define readability as "a quality that makes possible the recognition of the information content of material when represented by

alphanumeric characters in meaningful groupings, such as words, sentences, or continuous text." They qualify this definition by stating that readability is not based primarily on the attributes of the characters per se, but more on the spacing between words and sentences.

Display Luminance

Display luminance refers to the amount of light per unit area per unit solid angle leaving a surface (McCormick & Sanders, 1982). Most CRTs produce a maximum luminance of 68 cd/m² with some as high as 340 cd/m², while 65 cd/m² has been specified as adequate (Snyder & Taylor, 1979). Display luminance for dot matrix displays is typically 170 cd/m² (Riengold, 1974, cited by Snyder & Taylor, 1979). Throughout the research literature display luminance is typically reported as either character luminance or dot luminance and background luminance. Reflected luminance from the display is often not reported, although it adds to overall display luminance.

Shurtleff (1980) gave a recommendation of 34.3 cd/m² minimum for symbol luminances and states 68.5 cd/m² maximum is adequate for most applications. Recommendations for display luminance must be considered with contrast. If poor contrast exists, it is unlikely that a high character luminance will be appreciably better than a low character luminance. Gould (1968) pointed out that with low display luminances it is difficult to reduce the background luminance to maintain contrast. It must also be considered that for CRTs as luminance increases spot size tends to spread resulting in reduced sharpness of the image (Snyder & Maddox, 1978).

A Human Factors Society working group (1986) developing an American National Standard for visual display terminals (VDTs) recommends a minimum character luminance (or background, whichever is highest) of 35 cd/m².

For matrix-addressed displays, Snyder and Maddox (1978) recommended a dot luminance of ≥ 20 cd/m² (with dot modulation of 75%) for contextual displays, and a dot luminance of ≥ 30 cd/m² (with dot modulation 90%) for noncontextual displays.

Studies which have manipulated luminance in fact also manipulated contrast. There are no studies which have held contrast constant while varying luminance. Shurtleff (1980) stated that contrast may be low (2:1) when luminance is at 34.3 cd/m² or greater with a character size of 10 minutes of arc. However, when luminances are low the contrast ratio must be increased to a minimum of 5:1 and a visual angle of 20 minutes of arc. The studies which Shurtleff reviewed to make these recommendations did not use contrast ratios below 5:1 when investigating low luminance, and luminance level was confounded with contrast.

Table 11 lists recommendations for character luminance. Luminance appears to be a critical variable in user performance with matrix displays.

Luminance Contrast

Luminance contrast or (modulation) was defined in Section 2. Because many studies report only symbol luminance and background luminance equation (1) must be used to determine the contrast used in many studies; however, ambient illumination is also reflected off the screen. Therefore Gould's (1968) equation for defining contrast where reflected ambient illumination is considered is more appropriate. That is,

$$L = L_i + L_e, \text{ and } D = D_i + L_e, \tag{5}$$

where L_i is the internally produced symbol luminance, L_e is the luminance produced by the reflected ambient illumination, and D_i is the internally produced background luminance. Then

$$M = (L_i - D_i)/(L_i + D_i + 2L_e). \tag{6}$$

Howell and Kraft (1959) manipulated character size, contrast ratio (as defined by equation (1), and blur. Simulated CRT characters and numerals in the Mackworth font were used as stimuli. All main effects and the Contrast x Blur as well as the Contrast x Size x Blur interactions were significant. Figures 5 and 6 illustrate results for correct identifications and response speed, respectively, for the 86% and 95% modulation levels. There was little difference in performance when modulation was increased from 86% to 95% for characters larger than 16 min of arc. When characters were smaller than 16 min of arc or blurred, an increase in contrast was necessary. The authors recommend modulations of 94% with 88% considered acceptable. Gould (1968) stated that CRT displays typically have contrast ratios of 20:1 (90% modulation), but that this is hard to obtain without contrast enhancing devices.

Table 11
Character Luminance^a Recommendations for
Alphanumeric Legibility

Source	Recommendation
ANSI Draft, HFS-100	35 cd/m ²
Snyder and Maddox (1978)	Dot luminance ≥ 20 cd/m ² for contextual displays with modulation of 75%. Dot luminance ≥ 30 cd/m ² for noncontextual displays with modulation of 90%.
BRN	80 to 160 cd/m ²
Shurtleff (1980)	34.3 cd/m ² minimum; 68.5 cd/m ² adequate.
EG&G	65 cd/m ² minimum under sufficient contrast.
DCIEM	85 cd/m ² minimum.
VDT	45 cd/m ² minimum; 80 to 160 cd/m ² preferred.

^a Character or background luminance whichever is highest.

After reviewing a series of studies conducted in the Human Factors Laboratory at Virginia Polytechnic Institute and State University, Snyder and Maddox (1978) recommended a dot modulation (for matrix displays) of 90% for noncontextual displays and a dot modulation of 75% for contextual displays. Shurtleff (1980) recommended a modulation of 89% for characters smaller than 20 min of arc, and possibly higher yet for character sizes smaller than 10 min of arc.

The working group developing the ANSI VDT standard recommends a minimum modulation of 0.5 (contrast 3:1) with a modulation of 0.75 (contrast 7:1) being preferred. For characters smaller than 18 arcmin, higher contrast is required and may be calculated by:

$$\text{Luminance Modulation} = 0.3 + 0.07 \times (20 - S), \quad (7)$$

where S is the size of the characters in minutes of arc and luminance modulation is defined as in equation (1). Other recommendations are listed in Table 12.

Contrast is a critical display variable and has been found to interact with character size, ambient illumination, and many other variables. In general, when the display is degraded in some form such as small character sizes or high ambient illumination, a compensating larger contrast ratio or modulation is required to achieve a constant legibility.

Ambient Illumination

The effect of ambient illumination on displays is to reduce the displayed luminance contrast (Snyder & Maddox, 1978). Carel (1965, cited by Snyder & Maddox, 1978) illustrated that when ambient illumination at the display is 10 times greater than the display's background luminance, then the symbol-to-display-background contrast ratio must be significantly greater than when the ambient-to-display ratio is less than 10. There are not many studies which evaluated the effect of ambient illumination and its relationship to other display variables.

Burnette (1976) investigated the effects of ambient illumination, element (or dot) size, shape, and interelement spacing on a reading task and random and structured search tasks. Two levels of ambient illumination were evaluated: 700 and 5.4 lux. In general, the lower illumination level

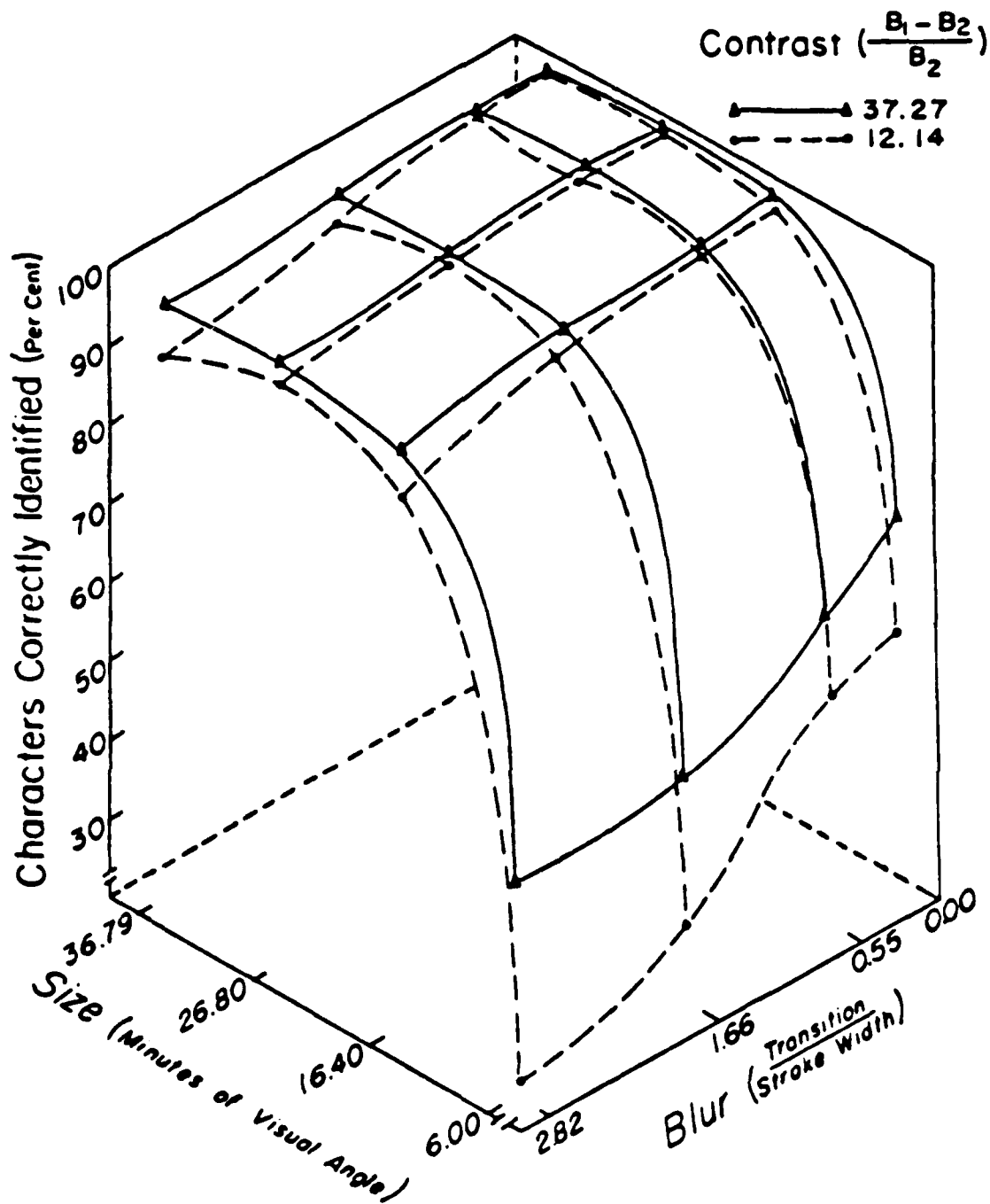


Figure 5. Percent of characters correctly identified as a function of size, blur, and contrast. Adapted from Howell and Kraft (1959).

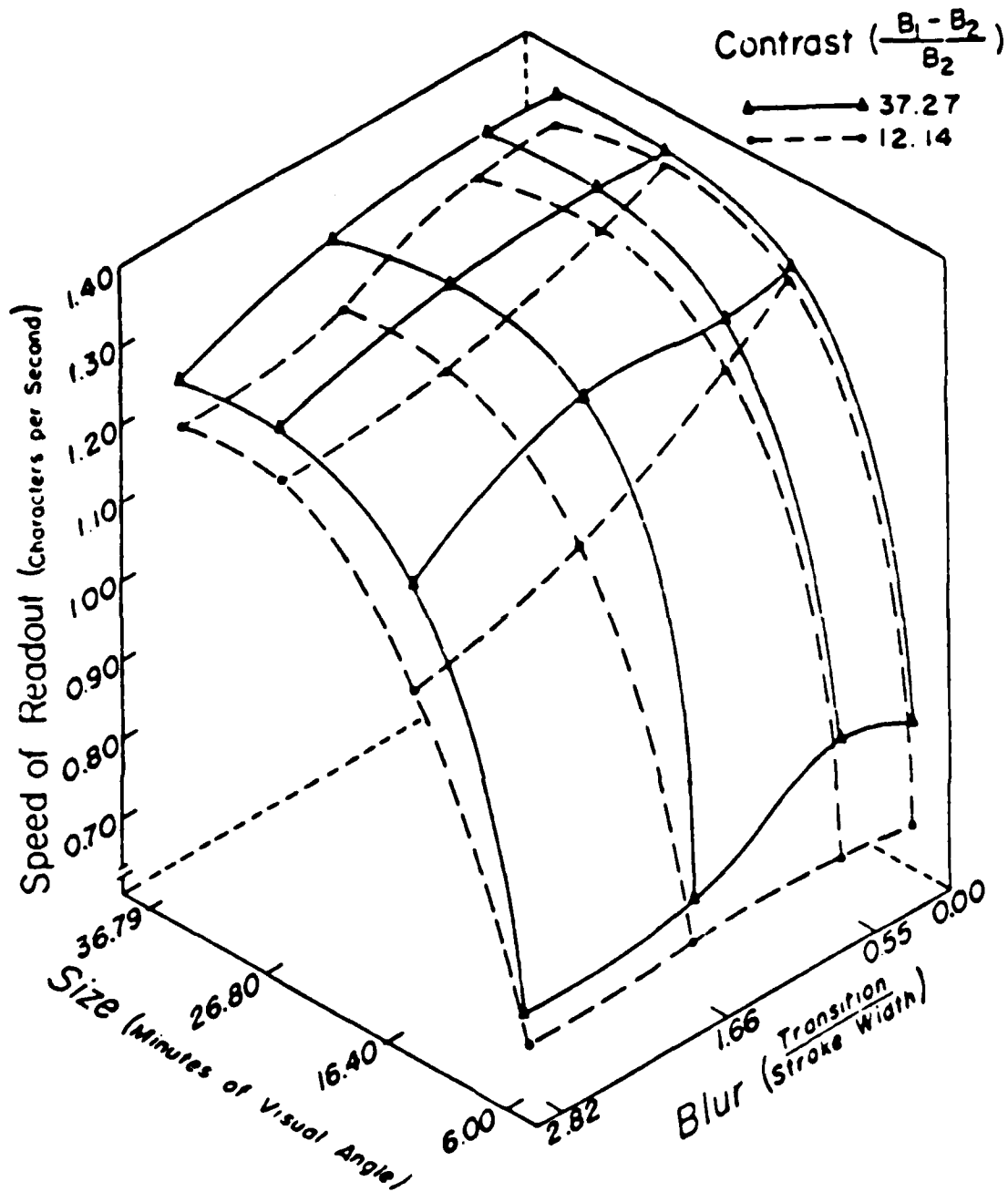


Figure 6. Speed of readout as a function of character size, blur, and contrast. Adapted from Howell and Kraft (1959).

Table 12

**Modulation Recommendations for
Alphanumeric Legibility**

Source	Recommendation
ANSI Draft, HFS-100	For characters larger than 18 arcmin, minimum modulation 0.5; 0.75 preferred. For characters smaller than 18 arcmin see formula 7.
Snyder and Maddox (1978)	Modulation 0.75 for contextual displays; 0.90 for noncontextual displays.
TUB	0.67 to 0.82 with background at least 20 cd/m ²
DIN	0.5 minimum; 0.71 to 0.82 preferred; 0.875 maximum.
IBM	0.875
MIL-STD- 1472B	0.82 minimum (white on black).
Shurtleff (1980)	0.89 for characters < 20 arcmin.
Howell and Kraft (1959)	0.88 minimum, 0.94 preferred for characters larger than 16 arcmin.
EG&G	Variable from 0.60 to 0.75 depending upon ambient illumination and user preference.
DCIEM	0.60 minimum in ambient of 750 to 1000 lux.
VDT	0.5 minimum; 0.78 to 0.82 optimum with background luminance between 15 and 20 cd/m ²

enhanced the modulation. Performance was superior with this level than with the higher illumination level.

Knowles and Wulfeck (1972) investigated varying levels of ambient illumination on four CRTs (three high contrast CRTs and one standard CRT). They were interested in determining whether "washout" occurred under high ambient illumination levels. The levels investigated were 1000, 10,000, 50,000, and 100,000 lux. Angle of incidence (30 and 60 degrees) and angle of regard (0 and -45 degrees) were also evaluated. The task used in this study was a discrimination task. Subjects were asked to indicate the location of a ring containing a 60-degree of arc gap. They performed this task while performing an auxiliary tracking task. Threshold detection data were collected. Results indicated that under the high ambient illumination of 100,000 lux none of the CRTs "washed out." In other words, all CRTs could be adjusted so that detection of the ring could occur even at 100,000 lux, under all viewing angles and angles of incidence. When the illumination angle of incidence was 30 deg, mean contrast values required for detection were of the same order of magnitude for both angles of regard (0 and -45 deg); however, when the observer position was -45 deg off-axis the mean contrasts required were 30% higher than for the 0-deg viewing position. In comparison, when illumination had a 60-deg angle of incidence there was a decrease in mean contrast required at the -45 deg viewing position for the three high contrast CRTs. An increase in mean contrast was required for the standard CRT. Unfortunately, this study did not indicate whether these differences were statistically significant, and the threshold data were only reported for the 100,000 lux illumination level.

Snyder and Maddox (1978) recommend an ambient illumination level of ≤ 125 lux for contextual displays and ≤ 75 lux for noncontextual displays.

The effect of ambient illumination for nonemissive displays or passive displays such as LCDs or ECs is another matter. These displays present information to the user by changing or modifying ambient illumination. An increase in ambient illumination in this case results in improved contrast.

Payne (1983) studied the effects of ambient illumination, angle of view, character subtense, and level of back light on an LCD. A central composite design was used. Illumination levels were 20, 390, 760, 1130, and 1500 lux. Subjects were asked to recognize four-digit numbers. Accuracy data were collected. The prediction equation resulting from the composite design indicated that error rates increased as back light and viewing angle increased and as character subtense and illumination decreased. The reliabilities of the partial regression coefficients for the independent variables were tested using an ANOVA. Viewing angle and ambient illumination were not significant predictors of error rate, while character size and back light were. Payne (1983) recommended maximizing ambient illumination levels and character subtense, and minimizing viewing angle and backlighting. With the central composite design, it is not possible to evaluate interactions between variables.

Duncan and Konz (1974) evaluated the effect of ambient illumination on the legibility of liquid crystal and light emitting diode displays. The display descriptions can be found in Table 13. Three levels of ambient light were investigated: 15, 150, and 450 lux. Subjects were asked to read digits on each display and data were collected for recognition time and the viewing distance at which no errors occurred. For the recognition time experiment, the digit size was held constant at 31 minutes of arc. Subjective measures of preferred illumination level, viewing distance, and display type were also used. Readers are referred to the study for the subjective results.

Table 13

Description of the Displays Used by Duncan and Konz (1974)

Display No.	Display Technology	Character Height	Percent Stroke Width-to-Height	Height-to-Width
1	7-segment LED	7 mm	2.6%	1.44
2	Hexadecimal LED	7 mm	4.7%	1.75
3	7-segment LED	19 mm	5.3%	1.58
4	3-1/2 decade transmissive LCD	11 mm	12.7%	1.69
5	3-1/2 decade reflective LCD	11 mm	12.7%	1.69

The results for recognition time are presented in Figure 7. The recognition time was longest for the transmissive LCD (4) at all three illumination levels, and was significantly longer than for all other displays. At the lowest illumination level (15 lux) recognition times using LEDs were significantly faster than recognition times using LCDs. This result is not surprising considering that LEDs are light emitting displays and LCDs are light modulating displays.

The LED displays 1 and 3 were not significantly different from one another and it appears that recognition time did not change as a function of illumination. On the other hand, the recognition times for the hexadecimal LED display (2) were significantly longer than the other LED displays,

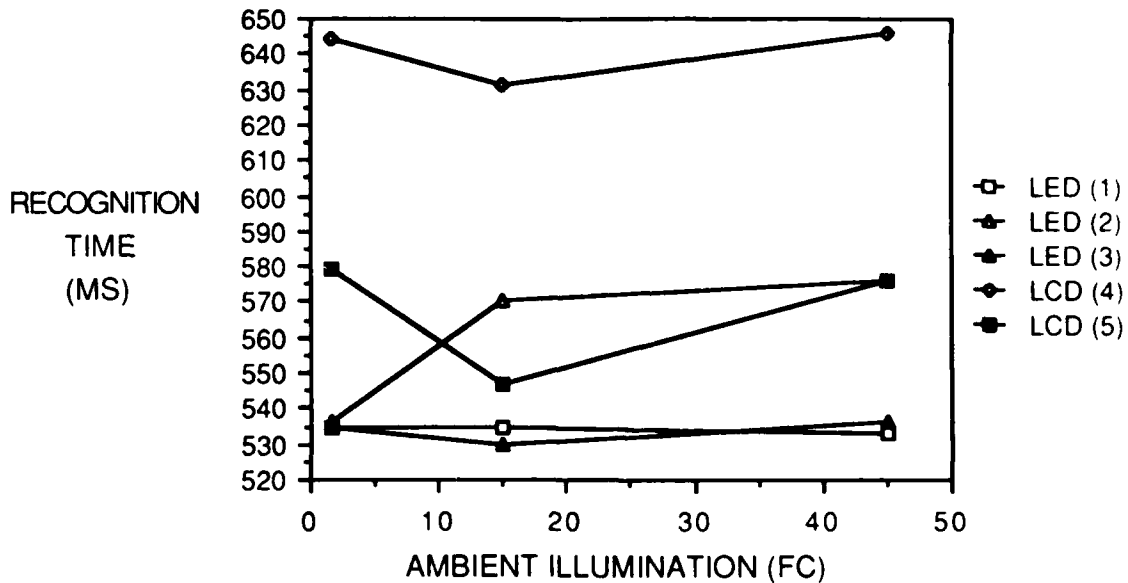


Figure 7. Mean values of recognition time for all digits at three levels of ambient illumination: From *Proceedings of the Human Factors Society 18th Annual Meeting*, 1974, p. 107. Copyright (1974), by the Human Factors Society, Inc. and reproduced by permission.

and as illumination increased recognition time increased. The authors did not discuss any possible reasons for the differential effects among LEDs. They did report the segment and background luminances for each display under all ambient light levels. From these data it is apparent that the hexadecimal LED (2) had a lower contrast ratio than the other LED displays and this could account for the results. The result of increased recognition time with higher illumination levels for this display is not surprising because, as previously pointed out, the effect of ambient illumination is to reduce the displayed luminance contrast for light emitting displays (Snyder and Maddox, 1978).

For the reflective LCD (5), as ambient illumination increased the luminance of the segment (or digit) increased resulting in a higher contrast ratio; thus, as would be expected, recognition time significantly decreased as ambient illumination increased. Recognition for the transmissive LCD (4) increased as ambient illumination increased from 150 to 450 lux.

No-error viewing distance results were influenced by character size with the largest viewing distance occurring for the largest character display (19 mm, LED 3). For this LED display (3) as ambient illumination increased from 150 to 450 lux, the no-error viewing distance decreased. For the reflective LCD (5), as ambient illumination increased the no-error viewing distance increased. The no-error viewing distances for the other displays were not differentially affected by illumination.

The studies by Payne (1983) and Duncan and Konz (1974) illustrate the effects of ambient illumination on passive or light modulating displays. However, the authors did not recommend ambient illumination levels for LC displays. Obviously, when considering what ambient illumination level is appropriate, the type of display used must be considered. Another consideration is that in many environments the ambient illumination is already fixed; therefore, it should be possible for display users to adjust contrast to compensate for inappropriate illumination levels. Table 14 lists recommended illumination levels for light emitting displays. No recommendations were found for light modulating displays.

Resolution

The term resolution is defined differently for discrete or fixed-element displays (such as flat-panel displays) and continuous displays such as CRTs. Lehrer (1985) and Snyder (1980) discuss

Table 14

**Ambient Illumination Recommendations for
Alphanumeric Legibility of Light Emitting Displays**

Source	Recommendation
ANSI Draft, HFS-100	200 to 500 lux
Snyder and Maddox (1978)	≤125 lux for contextual displays. ≤75 lux for noncontextual displays.
TUB	150 to 750 lux
DIN	300 to 500 lux for negative images 500 lux minimum for positive images
U of L	300 to 750 lux.
DCIEM	500 to 1000 lux
VDT	300 to 500 lux
SNBOSH	200 to 300 lux

the difficulty in defining the term resolution for CRT displays. This section will be divided into a discussion of literature pertaining to continuous displays and fixed element (discrete) displays.

Continuous Displays. There are many measures for CRT resolution in the literature. Resolution is often defined as the number of resolvable elements per unit dimension measured either subjectively or photometrically (Snyder, 1980). A common measure is the spatial frequency at which an observer cannot discriminate light and dark lines of an image. This measure is expressed in lines per unit display distance or per symbol height. Many studies have investigated the number of CRT raster lines per symbol height needed for optimum legibility. The general finding is that at least 10 lines per character height should be used (Buckler, 1977; Gould, 1968; Shurtleff, 1974; Winkler, 1979).

Erickson, Linton, and Hemingway (1968, cited by Snyder, 1980) found that recognition accuracy of alphanumerics improved as the number of lines per symbol height increased, or as the number of scan lines on the entire display increased. Shurtleff and Owen (1966b) compared 525 and 945 raster line displays. Alphanumerics were viewed at 6, 8, 10, and 12 lines per symbol height on each display. No significant differences in terms of response speed were found between the two systems. For correct identification the 525-line system resulted in poorer performance than the 925-line system at the 6 line per symbol height only. These findings conflict with Erickson et al. (1968).

Gould (1968) pointed out that character angular subtense interacts with number of scan lines, and therefore requirements for both parameters must be satisfied. He recommended a character size of between 12 and 15 minutes of arc and 10 lines per character height.

Discrete Element Displays. Resolution for discrete element displays is determined by the number of elements per display, element size, and interelement spacing. Density is a more appropriate term than resolution in this case (Snyder, 1980). Element shape is another consideration that has been found to have an effect on legibility. Actually, the research reviewed in this section was performed on raster CRTs, although the studies were simulating dot matrix characters that would be found on discrete element displays.

Resolution in the literature is often given in terms of the number of dots per unit area or dot matrix size. In a dot matrix display, element size, interelement spacing, and character size are necessarily confounded. If element size is increased, interelement spacing must decrease in order to keep the character size the same. Or if the spacing between elements is increased, the elements must be decreased in size. When researchers investigate the difference between 5 x 7 and other matrix sizes they are usually confounding interelement spacing, leaving character size and element size and

shape constant. Therefore, it is not possible to determine whether performance is a function of the number of dots in the matrix, the spacing between elements, or an interaction between them. The research on resolution for discrete element displays can be categorized into studies which investigate matrix size (confounding another parameter) and those which investigate the element size, shape, and spacing by holding matrix size constant.

Matrix Size

Matrix size, or the number of elements used to form an alphanumeric, has been investigated by several researchers; however, there has not been enough research to date to standardize selection of dot matrix size. A commonly used size is 5 x 7. A 5 x 7 matrix is made up of 35 elements, 5 columns of elements, and 7 rows of elements.

Shurtleff (1970a) was interested in determining the legibility of alphanumeric symbols formed from different matrices of dots. The dot matrix sizes investigated were 3 x 5, 5 x 7, 7 x 11, and 9 x 15. Character height, width, height-to-width ratio, stroke width, style, luminance, and luminance contrast were held constant. It is not possible to equate style (or font) exactly. However, Shurtleff tried to standardize the font by approximating the Lincoln/MITRE font as closely as possible for the different matrix sizes. The dot matrix characters were simulated using a CRT. Subtended visual angle was a group between-subjects variable for this study. Half the subjects saw the characters at 22 min of arc (Group A), and half at 6 min of arc (Group B, degraded viewing condition). Subjects were asked to read a 3 x 3 array of characters from left to right, top to bottom. Rate of correct identification (per minute) and percent errors were the response measures used. Two sessions were run per group to assess the effects of practice.

For Group A (22 minutes of arc) correct identifications per minute increased during both sessions as the matrix size increased from 3 x 5 to 5 x 7. A further increase in performance occurred for this group during the second session when the matrix was enlarged from 5 x 7 to 7 x 11. For the dependent variable percentage of errors there were no significant differences among the matrices for either session.

Results for correct identifications per minute for Group B (6 minutes of arc) indicate that there were no differences among matrices for the first session. However, there was a significant difference between 3 x 5 and 5 x 7 matrices for the second session. A main effect of matrix size was also found for percentage of errors. Post-hoc tests indicated a significant difference between the 3 x 5 and 5 x 7 matrices for the second session only: It was concluded that the 5 x 7 matrix is more legible than the 3 x 5 matrix. It was also concluded that a 5 x 7 matrix is just as legible as the larger matrix sizes used in this study, except that 7 x 11 is more legible when characters are large and the operator has practice.

The results of this study are rather surprising. It was expected that larger matrix sizes would be required for the degraded conditions based on the assumption that larger matrices make characters more legible. The author explains that there was an increase in performance from the 3 x 5 to the 5 x 7 in the degraded condition because the additional dots added to the 5 x 7 matrix added detail to the geometry of the characters. However, detail gained becomes less as even more dots are added; therefore, performance did not improve as matrix size was increased from 5 x 7. Other possible explanations are spurious effects, the small sample size, or between-subject variability.

Vartebedian (1971a) investigated the difference between 5 x 7 and 7 x 9 matrices and found the 7 x 9 matrix to be superior. This study is frequently cited in the literature as the basis for matrix size recommendations. Unfortunately, it does not appear that character size was the same for both matrices. The 5 x 7 characters were smaller; therefore, it is not surprising that the 7 x 9 characters resulted in better performance. Also, there were style differences between the character sets and possibly stroke width differences which confounded the variables and the results.

McTyre (1982) compared 7 x 7 and 7 x 9 matrices on two different CRTs. Unfortunately, dot size, upper case height, lower case height, and subtended visual angle were not held constant for the two different character sets. Even so, there were no significant differences between the two character sets nor between the two different CRTs.

The recommendations for using larger matrix sizes are based on the belief that the more dots per unit area the more similar the character becomes to the stroke character. This similarity is basically more a function of interelement spacing. The more elements per unit area the less space between elements, and symbols appear to be created from continuous strokes.

Snyder and Maddox (1978) investigated the effects of matrix size on the legibility of four different fonts. Three matrix sizes (5 x 7, 7 x 9, and 9 x 11) were investigated. Character size was allowed to increase as dots were added; however, they also designed 7 x 9 and 9 x 11 matrix size characters to remain the same size as the 5 x 7 characters by reducing the dot size and using the same

dot/space ratios for each matrix. This unconfounded the effects of character size and matrix size. Single characters were presented to subjects for recognition and error data were collected.

The main effect of matrix size was significant and results are illustrated in Figure 8. The 9 x 11 matrix with character size equal to the 5 x 7 character size resulted in the best performance in terms of recognition errors. The 5 x 7 matrix resulted in the poorest performance and was significantly different from all other character/matrix sizes. The 7 x 9 matrix resulted in the second poorest performance and was significantly different from all other character/matrix sizes. The 7 x 9 matrix equal to the 5 x 7 character size resulted in poorer performance than the 9 x 11 matrix and the 9 x 11 reduced character size matrix. The results of this study generally indicate that larger matrix sizes result in better single character recognition performance (fewer errors). Results also indicate that performance with the larger matrix sizes combined with the smaller character size was better than when a larger matrix with a larger character size was used. Because the dot/space ratios for the reduced character size matrices were the same, results cannot be attributed to higher percent active areas for these characters.

Recommendations for matrix size are listed in Table 15. Readers should consider that only the Snyder and Maddox (1978) study unconfounded the effects of character size, matrix size, and interelement spacing.

Table 15

Matrix Size Recommendations for
Alphanumeric Legibility

Source	Recommendation
ANSI Draft, HFS-100	5 x 7 numeric and upper case, 2 dots upward for diacritics. 7 x 9 for continuous reading tasks. Increase vertical height 2 dot positions upward for diacritic. For lower case, increase by 1 dot position downward, 2 or more positions preferred.
Snyder and Maddox (1978)	7 x 9 for contextual displays; 9 x 11 for noncontextual displays.
DIN	5 x 7 minimum. One additional dot position upward for upper case ascenders. Two dot positions downward for lower case descenders.
Shurtleff (1980)	7 x 9 or larger
DCIEM	5 x 7 minimum
VDT	5 x 7 minimum, 7 x 9 or greater preferred.

Element Shape, Element Size, and Interelement Spacing

Element shape. Vartebedian (1970a) evaluated circular versus elongated dots (elements) using a 7 x 9 dot matrix and concluded that the circular dot was superior to the elongated dot. The elongated dots used an increase in stroke width compared to the 5 x 7 dots, confounding the variables.

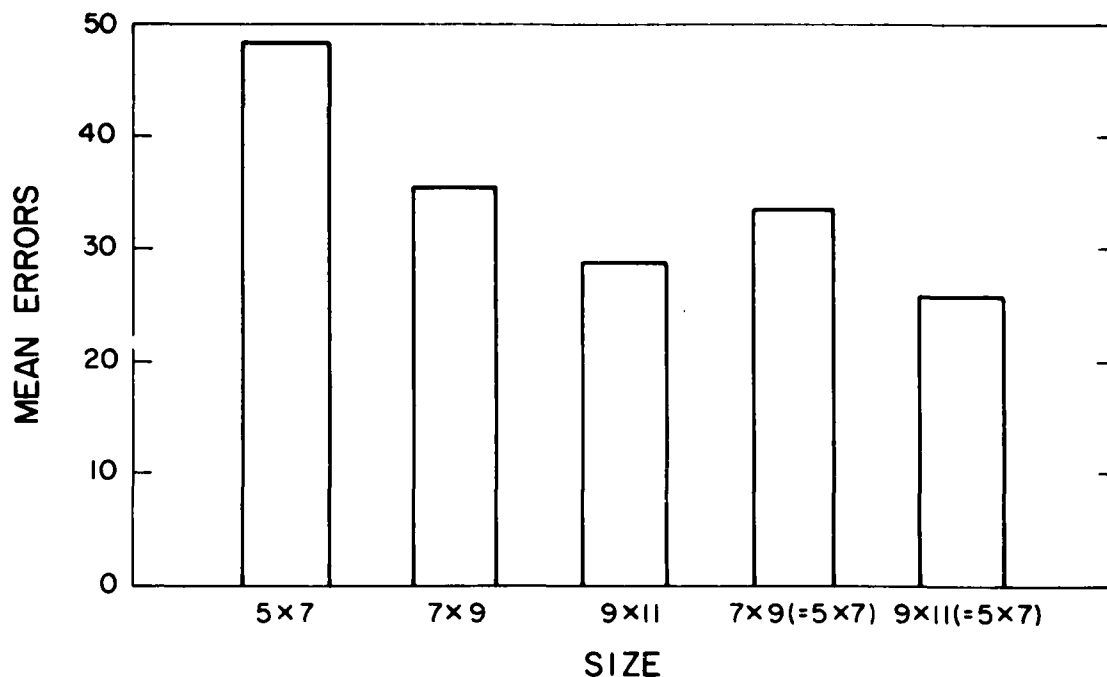


Figure 8. Effect of character/matrix size upon mean errors. Snyder and Maddox (1978).

Shurtleff (1980) believed that the number of dots in a matrix is of primary importance to identification and that element shape and interelement spacing are of secondary importance. He illustrated symbols formed by different numbers of elements (3 x 5, 5 x 7, 7 x 9, 9 x 11); however, spacing between dots was confounded in the examples as was stroke width, because when dots overlapped (due to smaller interelement spacing) stroke width increased (Sherr, 1979).

Williams (1981) reviewed dot matrix parameters in order to make recommendations for large screen displays. He stated that round or circular dots provide smoother characters and that square elements do not approximate stroke characters; therefore, they should be avoided except with large matrix sizes. He did not indicate any research sources to support this statement.

Semple et al. (1971) pointed out that round and square elements allow for presentation of alphanumeric and other symbols, lines, or shades of gray. Triangular or diamond shaped elements, however, may place restrictions on the character or symbol angles, causing them to lack smoothness. However, they believed that this may not be a problem when elements are sufficiently small with high density.

Element Size and Spacing. It is important to maximize the element size or area while minimizing the space between the elements (Semple et al., 1971). This result is typically referred to as the percent active area or fill factor and is defined as:

$$\text{Percent active area} = A/d^2 \times 100, \quad (8)$$

where A is element area and d is the distance or space between centers of two adjacent elements. Increasing element size or decreasing spacing between elements will increase percent active area (assuming character size is held constant), resulting in the extreme of characters which appear as if they were composed of continuous strokes rather than discrete elements.

Stein (1980) investigated the effects of percent active area on reading speed and accuracy. Full sets of alphanumeric characters varying in percent active area were presented to subjects who were asked to identify each character from left to right, top to bottom. Results of this study indicated that under ideal viewing conditions performance was unaffected by percent active area for active areas between 11.9% and 71.6%. When displays were degraded in some form, such as low luminance, low contrast, or small character sizes, a 30% active area was required to maintain performance for both reading speed and accuracy. Below 30% reading speed and errors increased, while above 30% there seemed to be little effect of active area on performance.

Vanderkolk (1976) studied two levels of 10 display parameters in a fractional factorial design. Percent active area was one of the parameters investigated, the two levels being 11% and 64%. Percent active area was highly significant ($p < .001$). Vanderkolk explains that the difference was due to perceived symbol brightness. As active area decreased the symbol brightness decreased. Percent active area was found to interact with several other variables, including surround luminance (0.17 and 342.6 cd/m^2), contrast ratio (0.5 and 3.0), and character subtense (15 and 30 minutes of arc). In all cases response time was significantly longer at the 11% active area for the two levels of the surround luminance, contrast, and subtense, with the greatest effect always occurring at the lower level of the particular variable. These findings again indicate that active area is important under degraded conditions.

Probably the most comprehensive study to date investigating element size, shape, and interelement spacing was conducted by Burnette (1976). In this study, square elements, vertical rectangular elements, and horizontal rectangular elements were simulated on a CRT (see Figure 9). Three levels of element size and three different edge-to-edge spacing ratios were also investigated under two levels of ambient illumination. Figure 10 illustrates the experimental design and the levels of each variable. The spacing between elements was different for each element size but spacing ratios were the same. A 5 x 7 matrix was used. All variables were treated as fixed-effects variables and factorially combined. Three different tasks were used: a reading test, a random search, and a menu search. Reading speed and average search times were measured.

For all three tasks performance was best with the square element. According to Snyder (1980), when the study was replicated using square and round elements the square elements were still superior for both reading speed and search. For the reading task the smaller element sizes resulted in faster reading speeds. Also, the closer together the elements were the faster subjects could read. For the search tasks the larger element sizes resulted in faster search times. Interelement spacing was not significant, nor was the interaction between element size and spacing.

Snyder (1980) discussed the findings of this study and explained that for reading tasks, smaller more compact characters minimize the number of eye fixations, thereby resulting in faster reading speeds. On the other hand, search tasks require peripheral detection; therefore, larger characters are required.

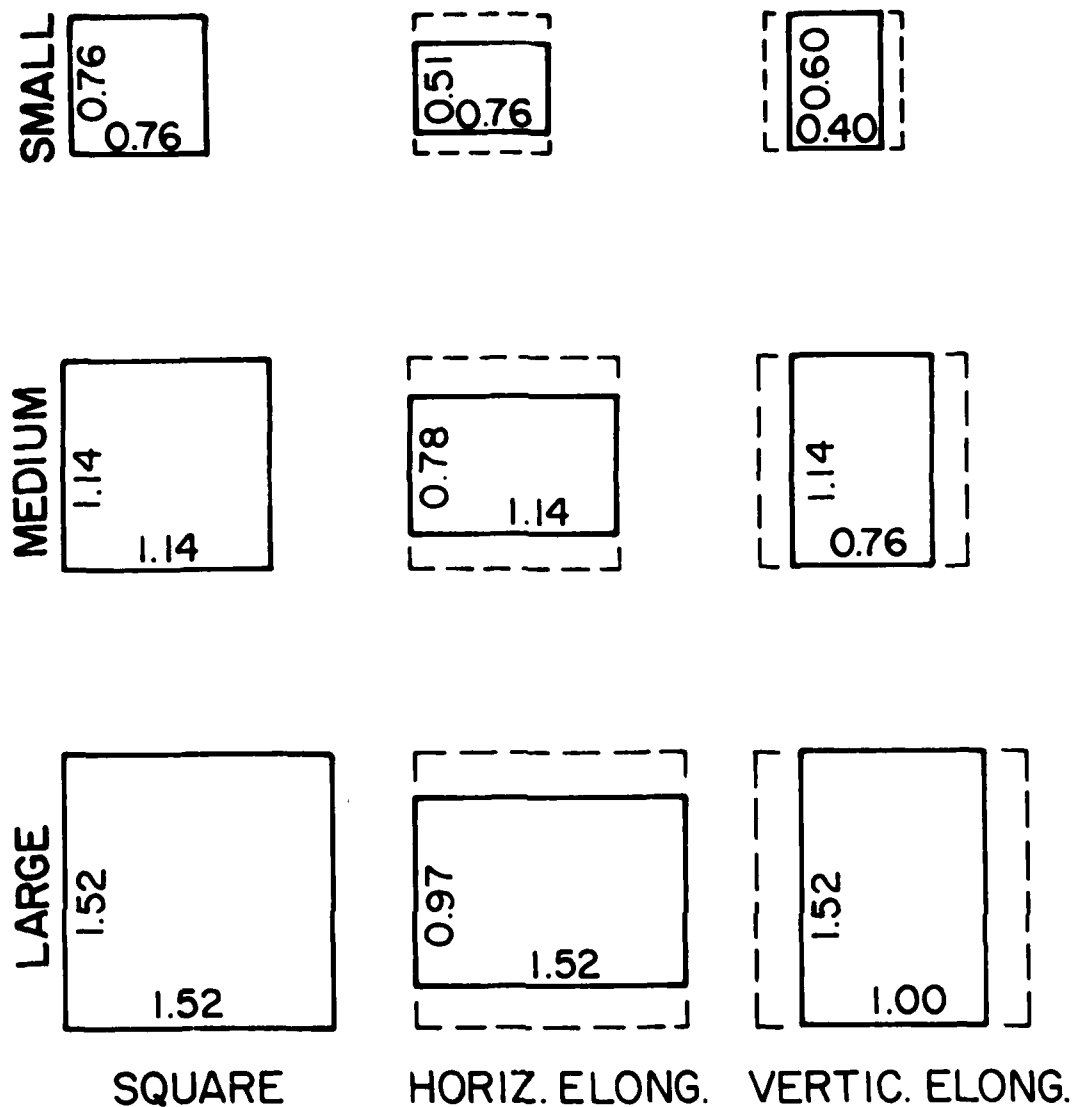
It should be noted that character size is necessarily confounded with element size and spacing. Also, this study simulated dot matrix characters and the luminance distribution across the individual elements was not uniform. The effect of this nonuniformity is unknown (Snyder, 1980). Because these data were for a 5 x 7 matrix size, the effects of the various parameters using other matrix sizes are similarly unknown.

Maddox (1977) performed a related experiment. In this study, three commercial dot matrix displays were simulated: the Burroughs Self-Scan IITM, the Owens-Illinois DIGIVUETM, and the prototype Westinghouse TFT (thin-film transistor) EL display. Three matrix sizes were used: 5 x 7, 7 x 9, and 9 x 11. Figure 11 illustrates the element sizes, shapes, and interelement spacings for the three displays. All displays were viewed under 5.4 lux. Also, the same font, as well as the same three tasks used in Burnette's study were used in this study.

Results for Tinker's Speed of Reading Test indicated a significant main effect of matrix size and interaction between matrix size and element shape. The 5 x 7 matrix resulted in significantly better performance than either the 7 x 9 or 9 x 11 matrices, indicating again that for reading tasks smaller characters are superior to larger characters. There were no significant differences in reading speed between the 7 x 9 and 9 x 11 matrix sizes. The interaction indicated that there were no differences among the three element shapes for the 5 x 7 matrix size. For the 7 x 9 matrix size, the TFT was significantly better than either the Self-ScanTM or the DIGIVUETM, and the Self-ScanTM was significantly better than the DIGIVUETM. For the 9 x 11 matrix, the DIGIVUETM was superior to both of the other two element shapes.

Results for the menu search task indicated a significant main effect of matrix size with the largest matrix size, 9 x 11 resulting in significantly faster search times than the 7 x 9 or the 5 x 7 matrices. Also, the 7 x 9 was significantly better than the 5 x 7 matrix. These results support the hypothesis that larger characters are required for search tasks. There were no significant effects due to element shape. Also, no significant effects were found for the random search task.

In summary, the literature dealing with resolution seems to indicate that when character size is held constant, enlarging matrix size improves performance. Characters made up of small dots and closely spaced elements are better for reading tasks, while larger characters are better for search tasks. Also, the results indicate that square elements are superior to circular and rectangular elements.



TYPICAL DIMENSIONS OF ELEMENTS (mm)

Figure 9. Element dimensions used by Burnette (1976).

Character Size

A great deal of research has been conducted to determine an optimum character size. Character size is typically specified by the subtended visual angle in minutes of arc:

$$\text{Visual } \angle \text{ (arcmin)} = (\text{character height} \times 3437.7) / (\text{Viewing distance}). \quad (9)$$

As previously reported, Howell and Kraft (1959) investigated the effects of character size, blur, and contrast on legibility of alphanumeric characters. Accuracy and response time data were collected. In general, it was found that 26.8 minutes of arc were necessary to maintain high accuracy performance under degraded conditions. An increase to 36.8 min of arc did not add to performance under degraded conditions. However, at 16.4 min of arc performance began to decrease under the highest blur and contrast conditions. Under the no blur condition, accuracy performance using 16.4

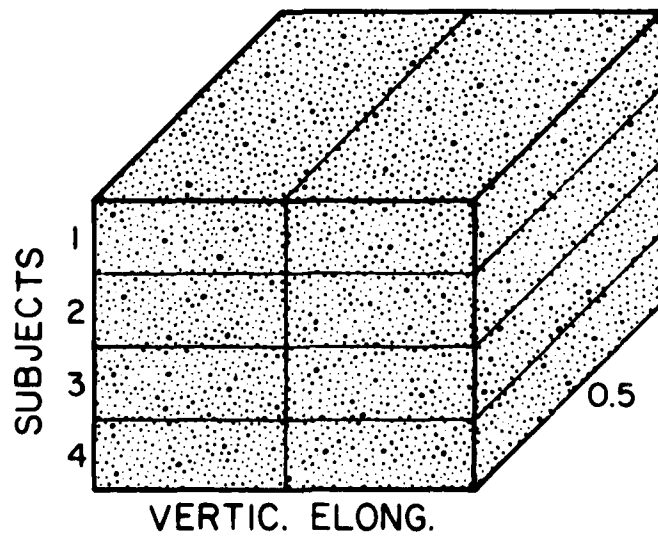
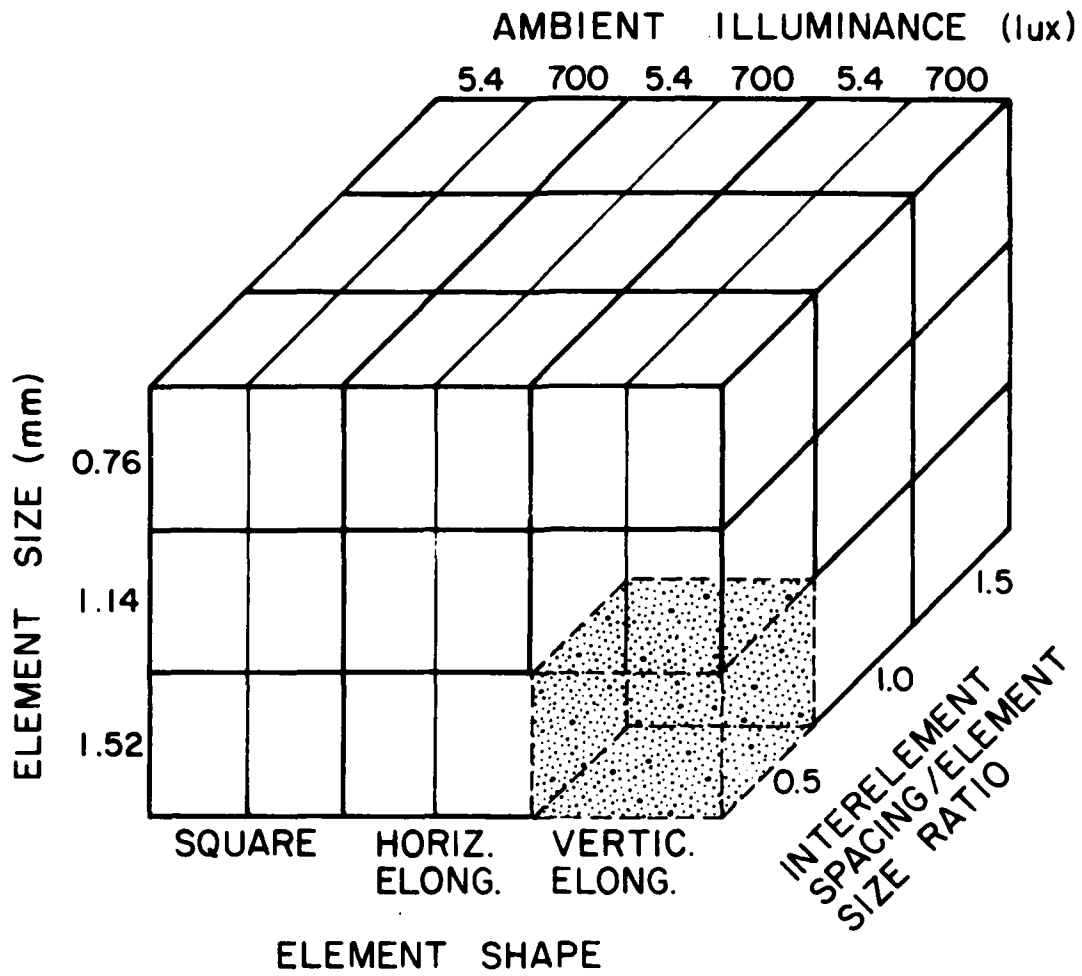
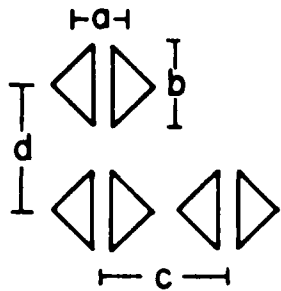


Figure 10. Experimental design used by Burnette (1976).

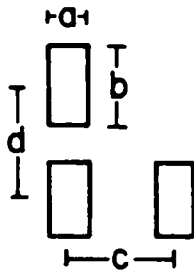
DIGIVUETM



$a = .813 \text{ mm}$
 $b = 1.240 \text{ mm}$
 $c = 1.697 \text{ mm}$
 $d = 1.697 \text{ mm}$

VIEWING DISTANCE =
203.2 cm

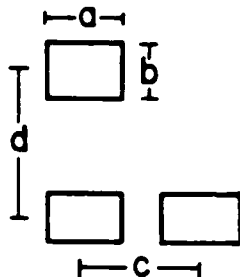
SELF SCANTM



$a = .610 \text{ mm}$
 $b = 1.067 \text{ mm}$
 $c = 1.554 \text{ mm}$
 $d = 1.554 \text{ mm}$

VIEWING DISTANCE =
101.6 cm

TFT



$a = .813 \text{ mm}$
 $b = 1.118 \text{ mm}$
 $c = 1.600 \text{ mm}$
 $d = 2.083 \text{ mm}$

VIEWING DISTANCE =
101.6 cm

Figure 11. Element sizes and shapes used by Maddox (1977).

min of arc was approximately equal to the performance for the larger sizes of 26.4 and 36.8 min of arc. For the response time data, there was similarly no difference between the two larger visual angles of 26.8 and 36.8 minutes; however, when the visual angle was decreased to 16.4 a performance decrement occurred. Therefore, it was recommended that when no blur exists a character size of 16 min of arc with a contrast of 37:1 (modulation of 95%) will provide 97% recognition accuracy. However, under degraded conditions of blur, the visual size should be increased to 26 min of arc.

Shurtleff, Marsetta, and Showman (1966) were interested in determining the visual sizes required to identify Leroy alphanumerics displayed at 10, 8, and 6 lines per symbol height. They found that for 85% identification accuracy a visual size of 7.58 minutes of arc was required when the characters were constructed with 10 and 8 lines per symbol height. For a 99% accuracy rate, a visual size of approximately 13 minutes of arc was required using 10 and 8 lines per symbol height. When the number of lines per symbol height decreased to 6, a visual size of 10.35 minutes was required for 85% accuracy, while a visual size of 35.97 was required for 99% accuracy using the standard Leroy font. (A revised Leroy font was also tested and results were very similar.)

Giddings (1972) investigated five alphanumeric character heights (0.25, 0.187, 0.156, 0.125, and 0.0625 inch) subtending 28, 21, 18, 14, and 7 min of arc, respectively. Characters were typed and a closed circuit television was used. Subjects were asked to read six-letter words and random digits, and reading speed and error data were collected. For reading speed performance, the main effect of character size was significant and there was an interaction between character size and type of material (words versus digits). Post-hoc analyses illustrated significant differences between words and digits for the character sizes subtending 14, 18, and 28 minutes of arc. A decrease in performance was found for both the smallest and largest character sizes. Giddings recommended an optimum character height of 0.156 inch for words and 0.187 inch for digits (18 and 21 min of arc).

Smith (1978) reviewed the literature to find the recommended standards for letter heights. He found that recommendations typically range from 10.31 to 24.06 minutes of arc with 5.16 minutes of arc the lower limit based on normal visual acuity. After determining what the recommendations were, a field study was conducted to find the legibility limit in angular subtense. It was found that a mean letter height of 6.53 minutes was the limit of legibility, while 10.31 minutes resulted in 90% legibility, and 24.06 minutes resulted in 100% legibility. The data were found to confirm many of the current standards for symbol size.

While investigating the sensitivity of response measures, Snyder and Taylor (1979) manipulated character size, display luminance, and viewing distance. Table 16 lists the character sizes in subtended visual angle for each of the seven viewing distances.

Table 16
Vertical Visual Angle Subtense (min of arc)
(Snyder & Taylor, 1979)

Character Size (mm)	Viewing Distance (m)						
	0.61	1.07	1.52	1.98	2.44	2.90	3.35
2.64	14.90	8.51	5.96	4.58	3.72	3.14	2.71
3.05	17.19	9.82	6.88	5.29	4.30	3.62	3.13
4.79	27.00	15.43	10.80	8.31	6.75	5.68	4.91
5.44	30.65	17.52	12.26	9.43	7.66	6.45	5.57

An analysis of variance was performed on accuracy and response time data. For accuracy data there was a significant improvement in performance as character size increased. Post-hoc comparisons indicated that the only single step improvement was between the 3.05 and 4.79 mm

character sizes. The interaction between display luminance and character size illustrated that the improvement of character size was greatest at 80 cd/m² followed by 27 cd/m², and finally 8 cd/m². It was also found that when viewing distance was increased (causing the subtended angle to decrease), performance accuracy decreased in general; however, the decrease was greater at the two smaller character sizes than the two larger character sizes. This effect was greatest at lower luminances.

The response time data showed significant main effects of character size, luminance, and viewing distance. As character size or luminance was increased, response time decreased. As distance increased, response time increased. An interaction between character size and distance was also found. As viewing distance increased, response time increased with the smaller characters, resulting in poorer performance than with the larger characters. Snyder and Taylor concluded that the legibility cutoff point for this study was for the character size of 4.79 mm viewed from a distance of 1.5 m, making the subtended visual angle 10.80 min of arc.

Character size is a critical design parameter in legibility. Character size has been found to interact with many variables. In general, under degraded conditions, such as low contrast and luminance, character size should be increased. Table 17 lists recommendations for character size. It is generally agreed that character size should be specified in angular subtense, not linear distance, units.

Table 17

Character Size Recommendations for Alphanumeric Displays

Source	Recommendation
ANSI Draft, HFS-100	16 arcmin minimum; 20-22 arcmin preferred No larger than 24 arcmin for reading tasks.
Snyder and Maddox (1978)	16-25 arcmin for contextual displays 1.0 - 1.2 arc degree for noncontextual displays
TUB	16 arcmin minimum; 20 arcmin preferred
DIN	18 arcmin for viewing distances > 50 cm;
Snyder and Taylor(1979)	10.80 arcmin
Howell and Kraft (1959)	16 arcmin (with modulation of 95%)
VDT	16-20 arcmin minimum
Giddings (1972)	18 arcmin for alpha characters; 21 arcmin for digits.

Stroke Width

Stroke width refers to the thickness of the stroke of the character and is generally used in conjunction with stroke-written as opposed to dot-matrix characters. A generally useful concept is the stroke width to character height ratio. According to McCormick and Sanders (1982), people can discriminate alphanumeric characters of a wide variety of stroke widths under nondegraded conditions. They recommend stroke width-to-height ratios of 1:6 to 1:8 for black characters on white backgrounds, and 1:8 to 1:10 for white characters on black backgrounds. Ratios for black characters on white backgrounds are lower than white characters on black backgrounds because white features appear to spread into adjacent black areas, whereas the reverse is not true (McCormick & Sanders, 1982).

Berger (1944, cited by McCormick, 1976) determined the average distance that subjects could read numerals of varying stroke width-to-height-ratios that were either black on white or white on black in daylight. The results have been plotted in Figure 12. The figure indicates that black letters have a lower optimum ratio (1:8) than white characters (1:13.3). In other words, a thinner stroke width is required for the white characters because of the phenomenon of visual spreading.

The recommendations above are for printed stroke characters. There has not been a great deal of research regarding stroke width for electronic display media. It is typically assumed that the recommendations for stroke width for printed material will hold for electronic displays.

In a review of the literature on the legibility of alphanumeric characters for electronic displays, Buckler (1977) recommended stroke width to height ratios ranging from 1:6 to 1:10 based on legibility for nonelectronic display media until data for electronic displays have been collected.

Crook, Hanson, and Weisz (1954, cited as reference 96 by Semple et al., 1971) investigated the effects of stroke width, contrast, character size, and symbol spacing on accuracy and rate of identification. A stroke width of 20% symbol height (or 1:5) was considered best. Stroke width did not affect accuracy when the modulation was above 90% and the characters subtended 22 minutes of arc. Stroke widths in this study were 9.8, 20 and 30% of symbol height. Stroke width also did not affect rate of identification when modulation was above 94% and characters subtended 22 minutes of arc.

In another study by Crook, Hanson, and Weisz (1954, cited as reference 95 by Semple et al., 1971) stroke width, symbol width, symbol spacing, and illumination were investigated. Figure 13 illustrates the levels of each variable investigated. All characters were 0.064 inch in height and subtended 15.71 minutes of arc. When the narrowest stroke width was viewed under the lowest illumination level, accuracy performance dropped. However, there were actually no statistically significant differences found for stroke widths for either correct number of responses or rate of identification. These two studies indicate that stroke width is an important variable when viewing displays under degraded conditions.

The working group developing ANSI standards recommends that stroke width or pixel dimension should be greater than 1/12 the character height. However, they also state that it is not an important variable in terms of performance when the character size, contrast, and luminance levels are adequate.

Defocusing of the CRT beam can cause variations in the stroke width; therefore, high contrasts should be used to minimize the effect of stroke width. High display luminance may cause stroke widths to vary as well.

Symbol Width

Symbol width refers to the width of the alphanumeric character. The ratio of the height to the width of the alphanumeric character is a typically more useful number. For printed text, McCormick and Sanders (1982) recommend a ratio of 1:1 for capital letters with a minimum of 5:3. For numerals 5:3 is recommended.

No data appear to exist for symbol width for electronic displays. Shurtleff recommended a ratio of at least 4:3 (or 75% of height) based on studies conducted by Crook, Hanson, and Weisz (1954, cited as reference 13 by Shurtleff, 1980), and Brown (1954, cited by Shurtleff, 1980). However, the study by Crook et al. only investigated two symbol widths, 86.3 and 59.8% of symbol height, while Brown investigated capital block letters used in aircraft plastic lighting plates under low luminance conditions (Semple et al., 1971). Therefore, recommendations from these data should be used cautiously.

The ANSI working group recommends ratios of 1:0.7 to 1:0.9 for column presentations. Element size and spacing will affect the height to width ratio.

Font

Font refers to the geometrical characteristics or style of the symbols or alphanumerics. Several researchers have been interested in determining optimum fonts for electronic displays. According to Sherr (1979), electronic displays are limited in the types of fonts which can be displayed based on the generation technique used, stroke or dot-matrix, with the dot matrix technique the most commonly used (Sherr, 1979). Fonts created on flat-panel displays limit font flexibility more than some cathode-ray tube devices (Abramson & Snyder, 1984).

A great deal of research exists regarding the legibility of stroke fonts (Cornog & Rose, 1967). Maddox, Burnette, and Gutmann (1977) point out that "it has not been satisfactorily demonstrated that the conclusions from stroke font research are directly transferable to dot-matrix fonts." Several researchers have been interested in comparing performance using stroke versus dot-matrix

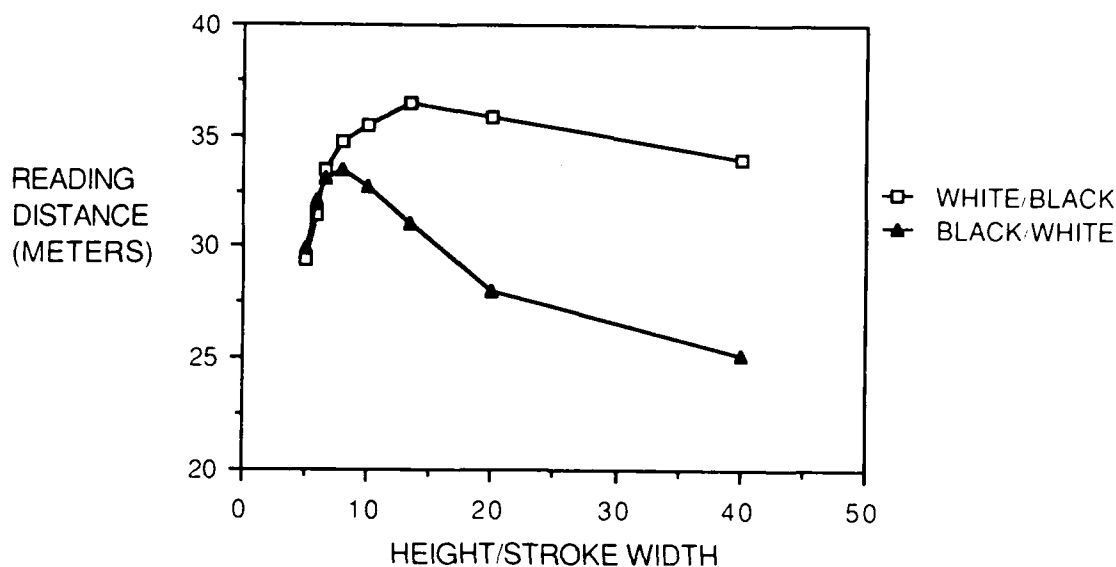


Figure 12. Average distance in meters at which black and white numerals with different stroke widths can be read. (Numerals were 42 by 80 mm). Adapted from Berger, 1944 cited by McCormick, 1976.

characters. If it can be demonstrated that there is no difference between the generation techniques, than stroke font research may be transferable to dot-matrix applications.

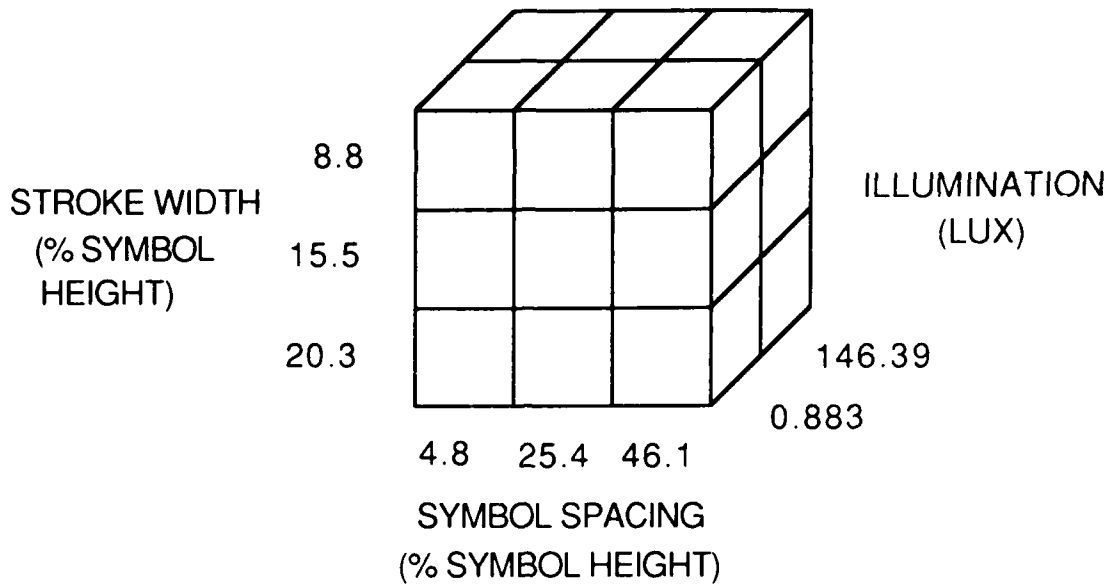
Dot Versus Stroke Characters. Vartebedian (1970a, 1971a) compared stroke versus 7 x 9 dot characters and stroke versus 5 x 7 characters. The stroke font was based on the Leroy font, whereas the dot fonts were designed by the author for maximum legibility. Response time and accuracy data were collected. He found no significant differences between the stroke and the 5 x 7 dot-matrix characters in terms of response time. However, the 7 x 9 dot font was significantly faster than the stroke font. There were significantly fewer errors using the 5 x 7 and the 7 x 9 fonts compared to the stroke font. Vartebedian concluded that dot-matrix generation is superior to stroke.

There are several problems with this study. First of all, there were character style differences other than those created by the generation technique. The stroke characters were based on the Leroy font; however, the dot characters were not designed to be as similar as possible to the Leroy font; therefore, font is confounded. Also, it appears that stroke width was not held constant. This is most obvious when comparing his 5 x 7 dot characters to the stroke characters.

In another study, Vartebedian (1971b) investigated generation method, letter size (.12, .14, .16 inch), and case (upper and lower). Again, the Leroy font was used for the stroke characters and a font designed by the author was used for the 7 x 9 dot characters, confounding font with generation method. Characters were presented on a CRT display. Subjects were required to search a display of 27 five-letter words to find a target word. The response measure was mean search time. The results indicated a main effect of case and subjects. (It is a questionable issue to test "subjects," a random variable!) There were no significant differences between generation methods, nor were there any significant interactions. Upper case words were recognized significantly faster than lower case words. Vartebedian, comparing the results with those of his previous study, states that single alphanumeric symbol legibility tests are more sensitive to generation method than a word search test due to the redundancy of the English language. However, single alphanumeric tasks are not representative of real world tasks. It is apparent that there is still not enough evidence to conclude that stroke research is directly transferable to dot-matrix fonts.

Font Comparisons. Considering that the dot-matrix technique is commonly used for display applications, there has not been a great deal of research comparing or developing fonts for dot-matrix displays, and there has been no standardized font for different matrix sizes (Maddox et al., 1977)

SYMBOL WIDTH: 59.8 (% SYMBOL HEIGHT)



SYMBOL WIDTH: 86.3 (% SYMBOL HEIGHT)

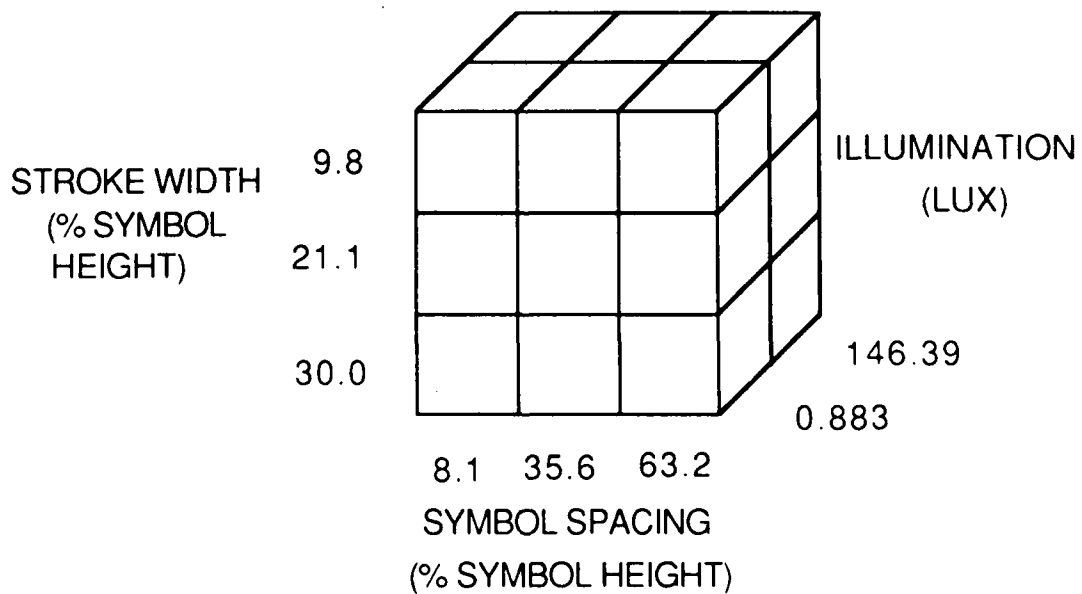


Figure 13. Experimental levels used by Crook, Hanson, and Weisz, 1954, cited by Semple et al. (1977).

A military standard font (MIL-M-18012B) was designed for aircrew displays (McCormick & Sanders, 1982). Ketchel and Jenney (1968) discussed the similarity of the Leroy font and the military standard, and state that based on evidence for the Leroy font the military standard is acceptable for electronic displays if departures are allowed due to the generation method used.

Vanderkolk, Herman, and Hershberger (1974, cited by Maddox et al., 1977) demonstrated that the dot-matrix adapted Lincoln/MITRE font is superior to other fonts.

Maddox, et al. (1977) compared three fonts in a 5 x 7 dot-matrix. A maximum dot font was created by using as many dots in the matrix as possible. A maximum angle font used fewer dots to give an angular appearance to the characters. These fonts were compared to the Lincoln/MITRE font used by Vanderkolk, Herman, and Hershberger (1974, cited by Maddox et al., 1977). Figures 14 through 16 illustrate these fonts. Single letters were tachistoscopically presented to subjects and accuracy data were collected. Significantly fewer errors were recorded for the maximum dot font than for either the maximum angle or the Lincoln/MITRE font. There was no difference between the maximum angle and the Lincoln/MITRE fonts. There was also a significant learning effect across trials; however, the differences among fonts remained the same. It should be noted that the maximum dot font had more dots, resulting in characters that appeared brighter although the dot luminance and size were constant across fonts. The percent active area and character sizes were the same for all fonts.

Snyder and Maddox (1978) performed a similar study which investigated three matrix sizes (5 x 7, 7 x 9, and 9 x 11) and four fonts (Lincoln/MITRE, Maximum Dot, Maximum Angle, and Huddleston). Accuracy data were collected. The character size was allowed to increase proportional to the number of dots in the matrix. They also created 7 x 9 and 9 x 11 characters keeping character size the same as the 5 x 7 size by proportionally reducing dot size and spacing. A main effect of font was found as was a font by matrix size interaction. Post-hoc comparisons revealed that the Huddleston and Lincoln/MITRE fonts were superior to the Maximum Dot and Maximum Angle fonts. There were no differences between the Huddleston and Lincoln/MITRE fonts. For the 5 x 7 matrix, the Huddleston font was superior, while for the 7 x 9 and 9 x 11 the Lincoln/MITRE font was superior followed by the Huddleston font. For the reduced 7 x 9 and 9 x 11 character sizes (each equal to the 5 x 7 in absolute size) the Lincoln/MITRE and Huddleston were not significantly different from each other and were superior to the other fonts. The authors recommended a choice between the Lincoln/MITRE and Huddleston fonts based on matrix size.

The studies comparing fonts used single-letter recognition tasks. Performance results may differ with reading tasks. While investigating the effects of dot and line failures on dot-matrix displays, Abramson and Snyder (1984) compared three fonts using a modification of Tinker's Speed of Reading Test. The three fonts investigated were: Huddleston, Lincoln/MITRE, and the font found on an HP2621A computer terminal (HP). The main effect of font was not significant. Complex interactions between font and the effects of percent failure, failure type (cell or line), and failure mode (on or off) were found. In general, the Huddleston font was found to result in the best performance, supporting the recommendation for the Huddleston font for maximum legibility and readability.

The study by Abramson and Snyder (1984) was the only study found which investigated the font found on a current production display. Further research which compares the fonts found on different display technologies has not been conducted.

Case. All the studies reviewed except that of Vartebedian (1971b) used only upper case alphanumerics. Vartebedian concluded that lower case words produced slower search speeds. Font design for lower case characters in matrix displays is difficult because of ascenders and descenders and fewer available dots. He also stated that a matrix larger than 5 x 7 is needed to provide legible lower case characters.

For continuous reading tasks, ANSI recommends a 7 x 9 matrix with two or more additional dot positions to accommodate descenders.

Abramson and Snyder (1984) compared line and dot cell failures on both upper and mixed case fonts. Their results indicated that when there were no dot or line cell failures, or when the dot and line cell failures were below 4%, there were no differences between upper and mixed case reading speeds. When failures increased to 8% or above, it took subjects significantly longer to read mixed case passages. As reported earlier this study used three different fonts. An interaction between font and case was not found.

A great deal more research is required to investigate user performance with lower or mixed

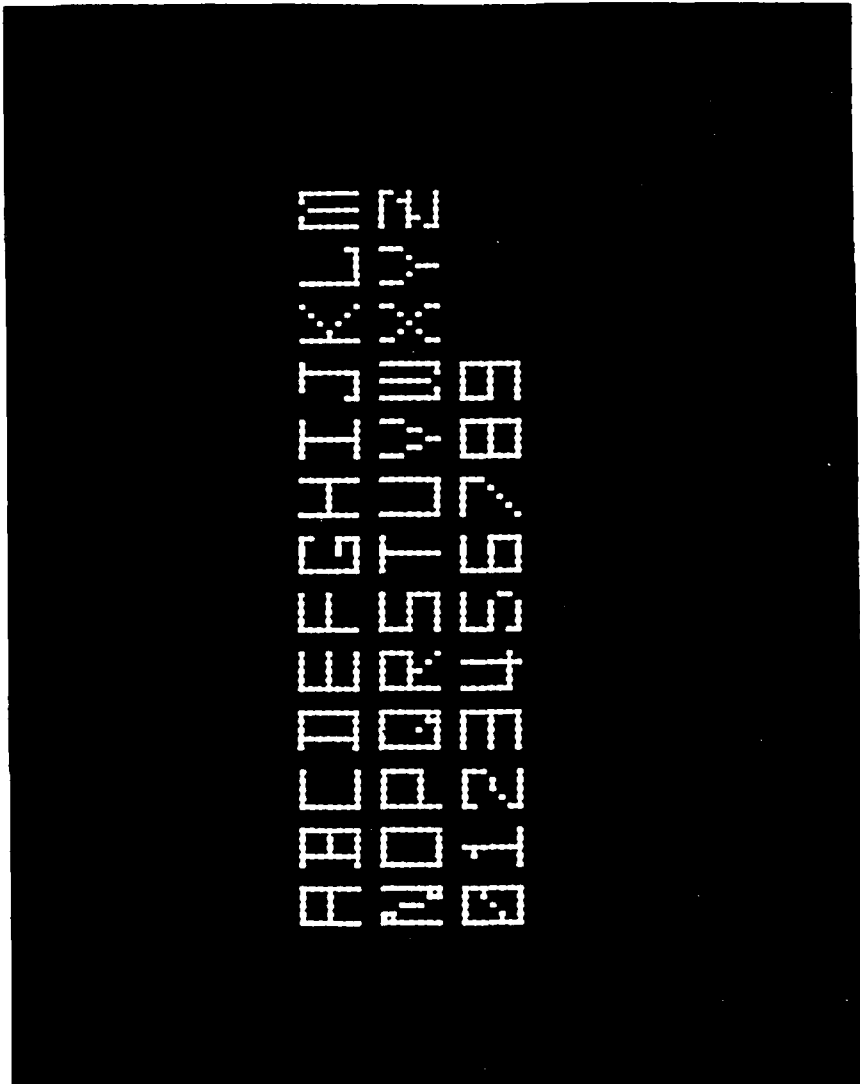


Figure 14. Maximum dot font in a 5 X 7 matrix used by Maddox et al. (1977).

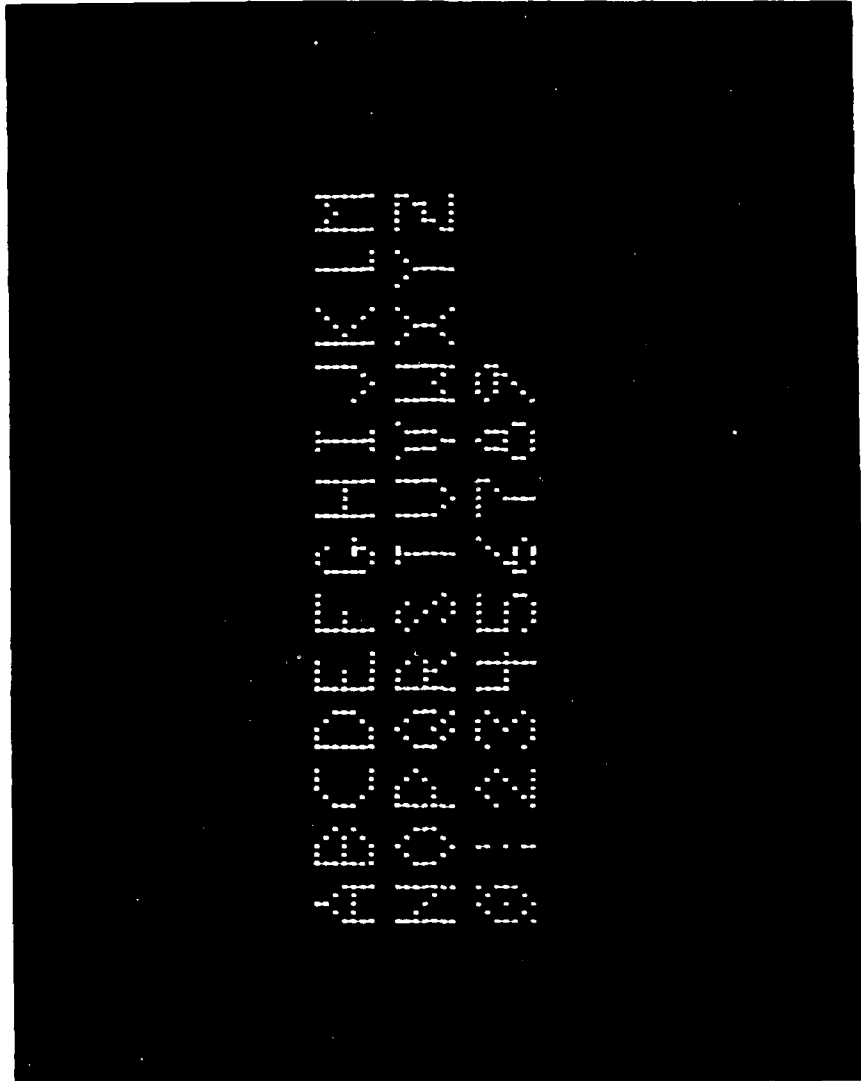


Figure 15. Maximum angle font in a 5 X 7 matrix used by Maddox et al. (1977).

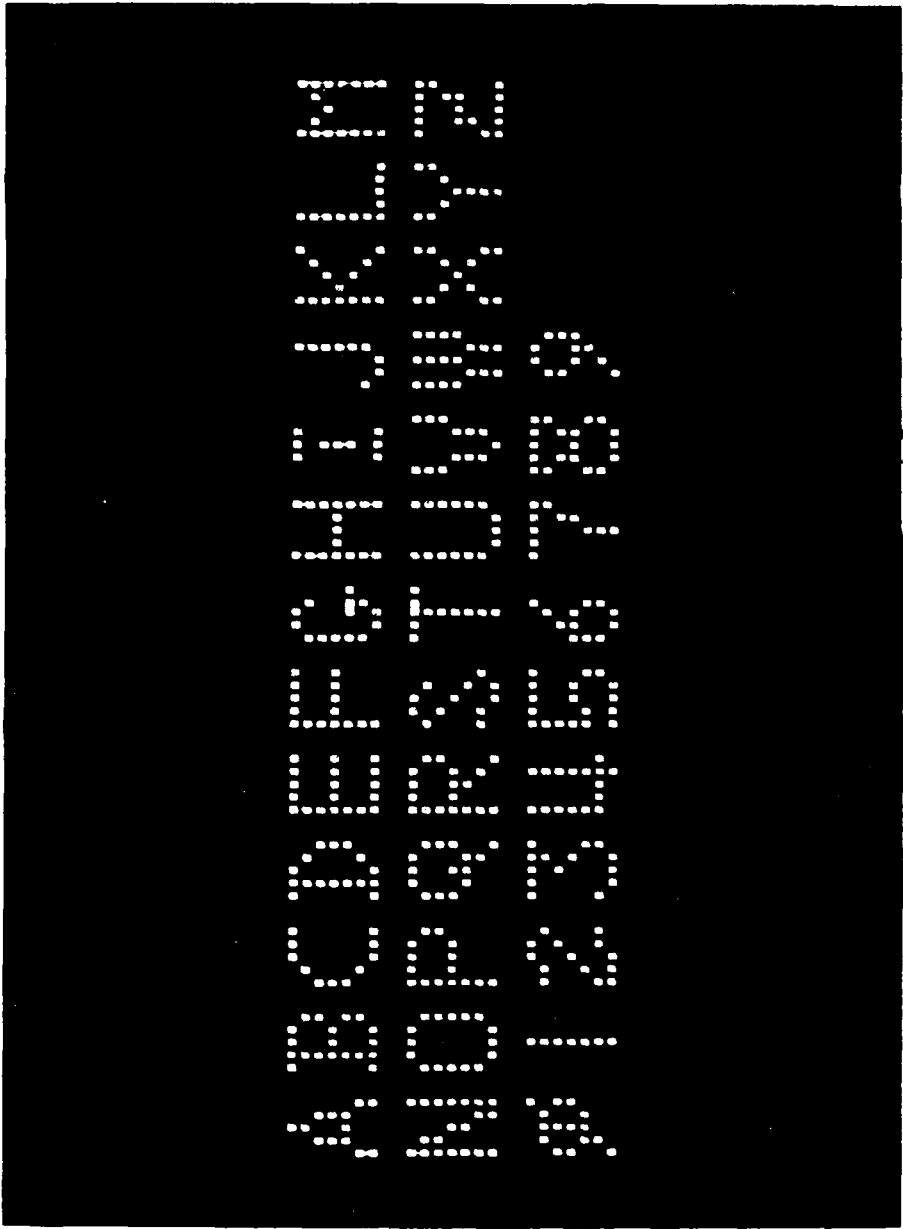


Figure 16. Lincoln/MITRE font in a 5 X 7 matrix used by Maddox et al. (1977).

case alphanumerics to determine the optimum lower case font.

Viewing Angle

Viewing angle refers to the angle between the viewer's line of sight and the display surface. Normal viewing occurs at 90 degrees. Often the luminance emitted from a display is directional; thus, as the viewing angle varies from normal, contrast is reduced. It was pointed out in Section 2 that contrast varies substantially with viewing angle for LCDs, which are nonemissive displays. The effect of viewing angle on performance will depend upon the display technology.

There is very little research regarding user performance and viewing angle for electronic displays. Seibert (1959, cited by Semple et al., 1971) found that alphanumerics on a television could be accurately identified at viewing angles from 90 degrees to 71 degrees. For viewing angles between 71 and 52 degrees, accuracy performance decreased. Semple et al. (1971) do not indicate the display parameters used in this study.

As mentioned in Section 3, Vanderkolk (1976) manipulated 10 parameters in a fractional factorial design. One of the parameters manipulated was viewing angle (90 and 45 degrees). There were no significant differences between the two viewing angles, nor were there any interactions with viewing angle and any of the other display variables. However, a full set of alphanumerics was not investigated in this study.

Snyder and Maddox (1978) investigated the effects of viewing angle (90 and 45 deg), and display type (DIGIVUE™, Self-Scan™ with round elements, and Self-Scan™ with square elements) on reading speed and visual search time. All displays resulted in significantly longer search times when the viewing angle was 45 deg versus the normal 90 deg angle. There was no effect of viewing angle on reading speed performance.

Reinwald (undated, cited by Shurtleff, 1980) conducted a study to evaluate viewing angle. He found that as the observer moved farther off-axis (deviated from a 90-deg viewing angle), the visual size of the symbols had to increase for performance to remain the same. He developed formulae to calculate effective viewing areas. In his review of this study, Shurtleff (1980) does not mention any of the experimental conditions or actual results of Reinwald's work. However, it appears that viewing distance, off-axis viewing, and character size were the only experimental variables. It is unlikely that these are the only parameters that would affect performance when viewing a display off-axis. Other parameters that may possibly affect performance include ambient illumination, curvature of the screen, contrast, resolution (Winkler, 1979), use of glare filters, display luminance, display size, and type of electronic display. Vanderkolk (1976) believes that stroke versus dot-matrix characters would not be differentially affected by viewing angle. However, this has yet to be confirmed.

Two important considerations for viewing angle are flicker and color. Peripheral vision is more sensitive to flicker than is foveal vision, especially at low illumination levels. Also, colors cannot be discriminated in the periphery because the color sensitive cones are found near the fovea. When viewing angles are other than the normal 90 deg viewers may be using some of their peripheral vision.

Character Orientation

Character orientation refers to the rotational positions of the character relative to the vertical display axis. There has been very little research investigating the effect of character orientation on performance.

Plath (1970) compared three sets of numerals: Air Material Equipment Laboratory (AMEL), slanted segmented numerals, and vertical segmented numerals. Five-digit numerals were presented to subjects using a slide projector, at three different presentation speeds (0.5, 0.1, and 0.02 s). Results indicated that the AMEL numerals were superior in terms of identification accuracy to either of the segmented numerals. No significant differences were found between the slanted and vertical segmented numerals. Readers should be cautioned that the stroke width of the AMEL numerals and the segmented numerals were quite different, and might have affected the results.

Vanderkolk (1976) found an interaction between character orientation (0 and 15 degrees) and character definition (7 versus 21 dots per symbol height). Response time was significantly slower for the 7 dots per symbol condition when characters were oriented 15 degrees. When characters were upright (0 degrees) there appears to be no difference between the two symbol definition conditions.

Vartebedian (1970a, 1971a) compared upright and slanted stroke characters and upright and slanted 7 x 9 elongated dot characters. There was no significant difference between the stroke upright and slanted in terms of response time; however, the stroke upright resulted in 2.2% fewer

errors, a statistically significant difference. The slanted 7 x 9 elongated dot characters resulted in significantly slower response times and significantly more errors (4.5% more errors) than the upright 7 x 9 elongated dot characters.

These studies seem to indicate that slanted characters may degrade performance on CRT displays. These studies used letter recognition tasks. Performance results may vary for word recognition or reading tasks. Also, it is possible that there may be differential effects depending on the display type or other display parameters. This parameter requires further investigation, particularly since rotating dot-matrix map displays appear to be likely in the next few years.

Temporal Characteristics

Isensee and Bennett (1983) investigated the perception of flicker and glare on a CRT display. The variables investigated were normal and reverse video, ambient illumination (100, 260, and 420 lux), and display luminance (120.1, 65.2, and 10.3 cd/m²). The off-axis angle at which flicker first became apparent was measured as the subject's chair was swiveled away from the face of the screen. Several results were found:

1. Flicker was perceived at smaller angles with lower levels of ambient illumination.
2. Flicker was perceived at smaller angles for the reverse video condition.
3. Smaller angles were reported as the display luminance increased with the smallest angles reported at the highest (120.1 cd/m²) display luminance. The main effect of display luminance accounted for most of the variance (61%).
4. An interaction between ambient illumination and display luminance was found. As display luminance increased, the effect of ambient illumination on the perception of flicker decreased. There were no differences between illumination levels for the highest display luminance condition (120 cd/m²).

The perceptual sensation of flicker is caused by the observer's ability to detect luminance changes when they are occurring at a rate below the integrating capability (time constant) of the eye (Sherr, 1979). Flicker is created on CRTs because the images are refreshed periodically by the electron beam. If the CRT is not refreshed frequently enough or if the phosphor does not have a long enough persistence, the display will flicker. In order for flicker not to be perceived, the displays must be refreshed above the observer's critical fusion frequency (CFF). The CFF is determined by requiring the observer to view an intermittent light. The intermittency rate is then increased until the observer sees only continuous light. That flicker speed is the CFF (Snyder, 1980). A large volume of data exists which discusses the effects of variables on CFF Brown (1965) and Sekuler, Tynan, and Kennedy (1981) provide reviews of this literature. Flicker has not been found to actually affect legibility, but it can cause observer fatigue and discomfort.

Snyder (1980) discusses the temporal contrast sensitivity function and its usefulness for predicting the frequency at which images will fuse on a display. While it would be difficult to determine the CFF empirically for all possible display conditions, an analytical prediction is quite feasible. Some of the display variables known to affect the CFF include phosphor characteristics, refresh rate, luminous intensity, screen size, rise and decay time, and ambient illumination (Semple et al., 1971; Snyder, 1980).

As defined in Section 2, rise time refers to the time period required by the device to reach maximum luminance after application of a square wave "on" pulse and decay time is the time it takes for the luminance to reach 10% of its maximum value after the pulse is turned off. If the decay times for a display are short, as with an LED, the eye does not have as much time over which to integrate the luminance as it would if the decay time were longer (considering the same luminous intensity in both cases). Therefore, the refresh rate must be greater for displays with short decay times so that flicker is not perceived. However, higher refresh rates generally produce higher luminance displays and, as Isensee and Bennett (1983) found, flicker is perceived more with higher display luminances.

The rise and decay times vary with the type of phosphor. Phosphors which have longer persistence (and therefore longer decay times) do not need to be refreshed as often as phosphors with shorter persistence. The goal is to limit the refresh rate so that the bandwidth necessary to carry the information is minimized. Display systems have a maximum bandwidth; therefore, increasing the refresh rate limits the information that can be transmitted to the display.

According to Snyder (1980), knowledge of a phosphor's Fourier fundamental modulation can be used with the temporal contrast sensitivity function to predict the refresh rate required to avoid

flicker. The Fourier fundamental modulation of the phosphor can be determined from the knowledge of the phosphor's persistence and applying the Fourier transform to determine the luminance modulation at the fundamental (refresh) frequency. Of course, some flat-panel displays (e.g., AC Plasma) avoid this problem entirely because they have inherent memory.

Uniformity

It is possible that typical levels of nonuniformity may have an effect on operator performance; however, there has been no research to indicate whether this is true. Large area nonuniformity may not be noticed by an observer if the changes are gradual; however, nonuniformity may become noticeable when the display is dimmed (Snyder, 1985). Farrell and Booth (1975, p. 3.2-60) quote technical reports which state that "a linear drop in luminance from center to edge of a rear projection display of two thirds was tolerable" and that "gradual brightness fall off of 50 percent will normally appear quite uniform". Unfortunately no performance data were given. The ANSI VDT standards working group recommended that the luminance on a display should not vary by more than 50% from the center to the edge or any other portion of the display.

There has also been no research regarding small area nonuniformity. According to Snyder (1985) small area nonuniformities can be predicted by comparing Fourier coefficients with the contrast sensitivity threshold function (CTF). If the coefficients exceed the CTF values, then observers will be able to detect the nonuniformity. There are no performance data to indicate an acceptable limit of (detectable or undetectable) small area nonuniformity.

Again, there are no performance data relating to edge discontinuities. Snyder (1985) stated that detection of edge discontinuities can also be predicted by comparing the CTF and Fourier analysis results. These parameters require further investigation.

Display Polarity

Display polarity refers to whether images on the display are light on a dark background (positive contrast) or dark on a light background (negative contrast). According to Rupp (1981) Europe is concerned with this topic and recommendations for positive image displays are typical. One concern is that when display users are refixating between a source document with dark characters on light backgrounds and a display screen with light characters on dark backgrounds, the pupillary response is taxed and may result in user visual fatigue. Rupp (1981) found that this was not a problem.

Bauer and Cavonius (1980) investigated the effect of contrast on the legibility of four-letter nonsense words. Polarity conditions were positive contrast with background luminance of 4 cd/m^2 (although the figure caption disagreed with the text by stating a background luminance of 10 cd/m^2), positive contrast with background luminance of 80 cd/m^2 , and negative contrast with background luminance of 80 cd/m^2 . Subjects were required to change their eye fixations from the screen to another display to simulate the situation where users are looking back and forth between the display and a source document. Error rates were collected. The authors equated stroke width by reducing the letters for positive contrast displays by 20% to adjust for the effects of irradiation or spread of light characters on a dark background (D. Bauer, personal communication, 1981). Results indicated that the negative contrast condition (at 80 cd/m^2) resulted in a significantly lower error rate than the positive contrast at (4 cd/m^2). The positive contrast (80 cd/m^2) condition was significantly worse than the other two conditions and observers complained that the letters were too bright.

In a review of the literature, Semple et al. (1971) found that display polarity did not have an impact on character identification. Shurtleff (1980) discusses two studies by Seibert. One study found that negative contrast was superior to positive contrast, and the other study found opposite results.

The ANSI working group states that either image polarity is acceptable as long as requirements for luminance, contrast, and resolution are met. They also state that dark characters on a light background may reduce distracting reflections from the display surface.

Isensee and Bennett (1983) found that flicker was perceived at smaller angles for negative contrast images. Therefore, a higher refresh rate may be required for displays with negative contrast (light background).

The results seem to indicate there is no legibility difference between positive and negative contrast displays. Whether or not there would be a differential effect due to dot or line failure has yet to be investigated.

Cell and Line Failures

Unique to matrix-addressed displays are the possibilities of individual line and cell failures (on or off). These failures have been found to reduce the legibility or readability of the display. Pastor and Uphaus (1982) point out three possible outcomes of cell or line failures: (1) the user can correctly identify the character or word; (2) the user is unable to identify the character or word; and (3) the user confuses the character or word with another. There are limited data to determine the amount of failure that is acceptable. These data are important to both display users and designers.

Riley and Barbato (1978) evaluated the legibility of five 5 x 7 dot-matrix alphanumeric fonts (ASCII, Lincoln/MITRE, Huddleston, ELLIS, and NAMEL) with discrete element degradation. In order to determine the importance of the element in a character, subjects were asked to identify dots in each character that would degrade the character the most if they were removed. Subjects were also required to remove dots so that the character was still easily distinguishable. This process allowed researchers to specify "importance values" for every dot. The same procedure was used to determine the effects of the addition of dots to a character on degradation. (All characters were presented on cardboard with black disks representing the dots in a character.) After determining the "importance values" of each dot, each character was degraded by either removal of dots, addition of dots, or the simultaneous addition and removal of dots. No differences among the fonts were found and neither the removal of dots, addition of dots, nor the simultaneous addition and removal of dots differentially affected character identification. It should be considered that for electronic dot-matrix displays, dot (or line) failures, either off or on, are random.

Pastor and Uphaus (1982) evaluated 7 x 9 ASCII numbers for confusability with other ASCII numbers under two percent dot loss. Results indicated that a linear relationship exists between specific dot loss and identification accuracy.

Spencer, Reynolds, and Coe (1977, cited by Abramson & Snyder, 1984) found that readability decreased for four different typefaces as background noise levels consisting of random dots increased.

Laycock (1985b) developed a procedure on an image processor to store and systematically add failures to text images. Laycock subjectively evaluated the failures and made several conclusions. He determined that cell failures which were failed "on" were more disturbing than "off" cell failures and less than 0.01% of "on" cell failures is tolerable while 1.0% of "off" cell failures can be tolerated. He also concluded that a single line failure may be unacceptable if it aligns with major components of text characters. For cell failures, lower case text degrades more rapidly than upper case while line failures have approximately the same effect on both cases. Line failures were believed to affect text with no character redundancy (e.g., abbreviations and mathematical formulae) more than text with redundancy. The author points out that the conclusions were subjective opinions made by the author and that a statistically valid study is necessary.

Abramson and Snyder (1984) had previously conducted such an investigation. They evaluated the effects of cell and line failures on readability of an AC plasma display. The parameters investigated were font, case, failure mode, failure type, and the percentage of cells failed. Figure 17 illustrates the levels of each variable and the experimental design. A modification of Tinker's Speed of Reading Test was used. Response time and the frequencies of correct, incorrect, and null responses were collected. The effects and interactions of these variables were complex. In general, the results indicated the following:

1. Random cell failures, either off or on, resulted in the longest reading speeds and the most incorrect and null errors.
2. Off failures generally resulted in better performance than on failures but this was dependent upon failure type. When failures were off, line failures resulted in poorer performance than cell failures. However, when failures were on, cell failures resulted in poorer performance than line failures.
3. Upper case presentations resulted in significantly better performance than mixed case under all conditions.
4. No main effect of font was found for either reading speed or response frequency; however, many interactions were found. In general, the Huddleston font was found to be the most resistant to degradation.
5. As the percent of failures increased above 2%, response time and the number of incorrect and null responses increased. These results indicate that if the failure rate is kept at 2% or below, degradation has a minimal impact on performance.

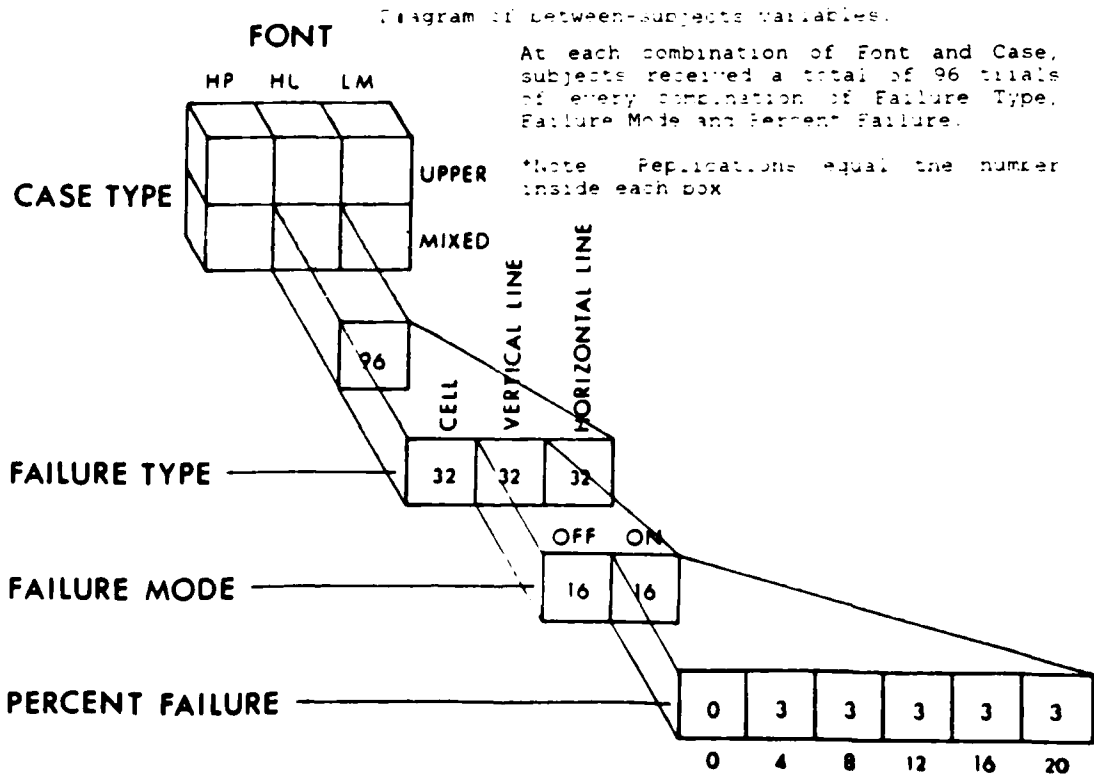
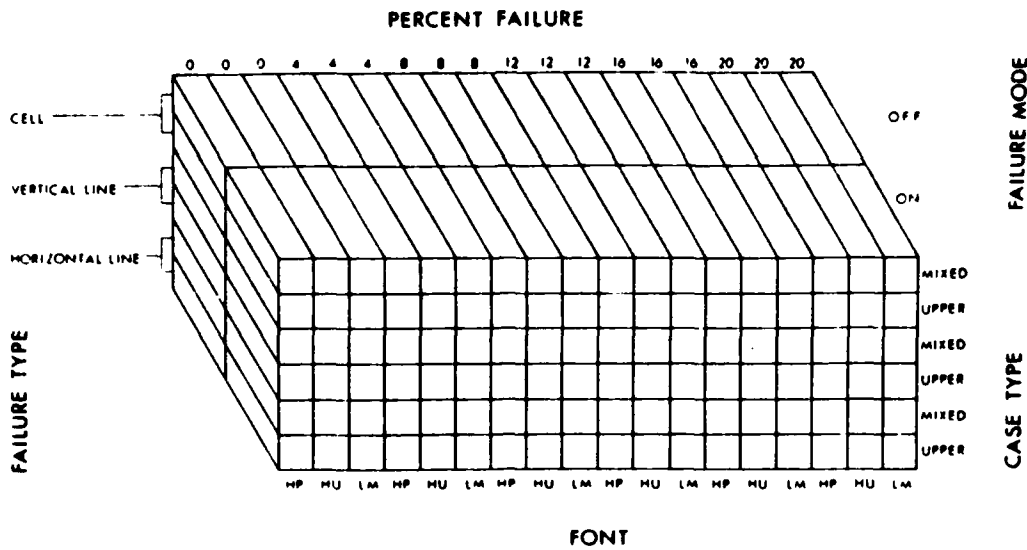


Figure 17. Experimental design used by Abramson and Snyder (1984).

The effect of display failure on human performance is complex. Many variables were found to interact and influence readability. It is likely that other display variables will also interact, such as display polarity, contrast, and character definition parameters. A great deal more research is required to make recommendations for acceptable limits of dot and line cell failures on alphanumeric displays.

Chrominance

The research discussed in this section has been performed using monochromatic displays, typically achromatic. Many flat-panel technologies are only available in a given wavelength, although full color is quickly becoming available. There is little information on how the use of color affects legibility, although color is being used to code information on displays under the assumption that it may enhance performance. With the advent of computer graphics technology and the availability of full color flat-panel displays the need for criteria for using color is essential.

A great deal of human performance research with color has been conducted. Krebs, Wolf, and Sandvig (1978, cited by Snyder, 1980) reviewed and analyzed the color literature. Wagner (1977) prepared an annotated bibliography of studies which investigated the use of color on television displays. In general, the research points out that the effect of using color depends on the specific application. General rules are often expressed, such as "untrained observers can only discriminate up to nine colors adequately" (McCormick & Sanders, 1982); "selected colors should be widely spaced in wavelength" (Krebs et al., 1978 cited by Snyder, 1980); or "blue leads to poor legibility" (Myers, 1967, cited by Snyder, 1980). For the most part the researchers perform no radiometric measurements for specifying the color of the stimuli. Colors are typically described by their subjective labels (e.g., blue). Also, many of the studies have investigated the use of color stimuli on black or achromatic backgrounds. Subsequently "quantitative criteria for color coding and for estimating the efficacy of color coding are essentially non-existent" (Snyder, 1980).

Reviewing all of this literature again in this report would lead to the same conclusions with little information that could be applied accurately to display design. Therefore, it seems more appropriate to discuss some recent research which has been concentrating on developing quantitative metrics for predicting performance with color. Also, some of the perceptual problems with viewing colored stimuli will be reviewed briefly.

Color Contrast. The importance of adequate contrast for legibility has been pointed out in previous sections. Contrast in the studies discussed thus far was a measure of luminance contrast. With this measure of contrast, human visual performance can be predicted (Snyder, 1980). However, visual performance cannot always be predicted for stimuli of one color against a background of another color because an adequate measure of color contrast has not existed until recently.

Most recently, several researchers at Virginia Polytechnic Institute and State University have tried to develop a measure of color contrast that can be related to human visual performance. These studies determine the color difference (in linear distance) between a target's color coordinates and its background color coordinates within a given color space. In order for a linear color distance to be obtained it is necessary to have a color space that is perceptually uniform. The original 1931 CIE tristimulus space was found not to be perceptually uniform; that is, equal distances on the color diagram did not correspond to equal perceptions (Post, 1983). Since 1980, considerable research has been conducted trying to develop a uniform color space. The color difference within a uniform color space is used to represent the magnitude of color contrast. The measure of color contrast (or difference) is then correlated with human performance (e.g., reading speed).

Carter (1982) has used CIE $L^*u^*v^*$ color difference formulae to come up with an algorithm to determine the best set of CRT display colors based on the number of colors needed, their chromaticity coordinates, the luminance range of the phosphor, and the number of equal luminance steps of the phosphor. The algorithm outputs a set of N high contrast colors. De Corte (1985) adapted the algorithm to take ambient illumination into consideration. Ambient illumination has been demonstrated to affect legibility performance with color displays (Ellis, Burrell, Wharf, & Hawkins, 1975; Snyder, 1980).

Post, Costanza, and Lippert (1982) compared three uniform color spaces--1976 CIE $L^*u^*v^*$ (Luv), 1976 CIE $L^*a^*b^*$ (Lab), and Cohen and Frieden's Wab--and developed equations to transform color differences in each space into equivalent achromatic contrast. It was hypothesized that if the color contrast could be transformed to achromatic contrast, then the knowledge already obtained about achromatic displays could be directly applied to color displays. The color differences in each color space were regressed on achromatic contrast settings that were obtained by having subjects adjust the contrast on a achromatic pair of stimuli to match the color contrast of an chromatic pair of stimuli. The color pairs had previously been matched in (subjective) brightness. Single factor linear regression indicated that the three color spaces were not uniform for predicting the achromatic contrast. A three-factor second order linear regression was then performed with one factor for each of the axes in the color space. The results indicated that, for the two CIE spaces, distances along the L^* axis contributed more to the equation than did the chromatic axes, substantiating the belief that the color spaces were not uniform. The Wab space, on the other hand, appeared uniform for predicting achromatic contrast. Results with the Wab space also

indicated that the color difference alone may not be adequate for representing color contrast. Therefore, a new metric was formulated and it was regressed on the achromatic contrast instead of the color difference in Wab space. The new metric was

$$C_{\text{mod}} = dc / (R_1 + R_2). \quad (10)$$

where dc is the color difference in Wab space and R_1 and R_2 are the distances of each color from the origin of the color space (black). This metric was found to be a better representation of color contrast in Wab space than were the color distances alone.

To see if the results generalized, another experiment was conducted. This time the stimuli (color pairs) also varied in brightness as well as in hue and saturation. The same regressions were calculated. Results indicated no significant differences between the three-factor Luv and Lab models. However, Luv and Lab model regression coefficients were significantly greater than the C_{mod} -Wab regression coefficients. Results of the two studies were compared by using the regression coefficients of the second study to predict the first and vice versa. In both cases, the Lab coefficients predicted the results in either direction. The authors concluded that although the two CIE color spaces are not uniform they may be used for specifying color contrast, but the axes should be rescaled.

Lippert, Farley, Post, and Snyder (1983) performed a related study in which color differences in three color spaces (Luv, Lab, and Wab) were regressed on the dependent measure of response speed. In this study targets (3, 4, or 5 digit string of dot-matrix numerals) of three different colors (achromatic, yellow-green, and red) were presented against eight different background colors in a darkened room. The target luminances were held constant at 46.6 cd/m^2 while each of the background colors were presented in seven different luminances. The luminance modulation " L_{mod} " for each target background combination was calculated using the equation

$$L_{\text{mod}} = (LT - LB) / (LT + LB), \quad (11)$$

where LT is target luminance and LB is background luminance. Targets were presented to subjects who read the numerals and response time data were collected and transformed into response speed (responses per second).

For all three target colors there were no significant differences in performance among all background colors at the two highest L_{mod} s, 0.270 and 0.316. At lower L_{mod} levels, red and purple backgrounds resulted in the best performance for all three target colors. Red targets resulted in faster reading speeds than achromatic or yellow-green targets for all backgrounds except purple. The color differences between target and background were calculated in $Y_u'v'$, Lab, and Wab color spaces. $Y_u'v'$ is a color space which utilizes the CIE $u'v'$ coordinates and the target and background luminance difference (Y). These color differences (in linear distance) and a term for length of target string were regressed on response speed in two- and four-factor linear regressions. Reading speed could be predicted by color difference depending on the color space used. The $Y_u'v'$ model provided the best results ($R^2 = 0.755$).

Post (1983) performed a similar study; however, the models were developed to predict response speed from color contrast for reading dot-matrix numerals presented against digitized full-color photographic backgrounds. In this study, five different target colors were used (red, blue, green, yellow-green, and achromatic). Response speed was best for the red target, followed by blue, green, yellow-green, and achromatic. Post-hoc comparisons showed that response speed for red and blue were significantly faster than for the other three colors. These results are interesting in that other researchers have recommended that blue leads to poor legibility (Krebs, 1978; Myers, 1969 cited by Snyder, 1980)]. Post stated that the difference may be that achromatic backgrounds don't generalize to colored backgrounds.

Post also performed regression tests to determine the color difference between targets and backgrounds in Luv, Lab, Wab, and the traditional CIE Tristimulus space (Tri) as a control. It was not practical to determine color contrast between the target and every point in the background cluster close to the target. Therefore, an "average color" of the background in each color space was determined by averaging over all the background pixels within a 2-degree radius of the center of the target. Two-factor and four-factor regressions on reading speed performance were performed. The two-factor regression showed no practical differences among the color spaces. Four-factor regressions (requiring linear rescaling) produced the best results for Luv and Lab color spaces ($R^2 = .480$ and $R^2 = .496$, respectively), which indicated that simple linear rescaling improves their perceptual uniformity. This was not the case with Wab or Tri color spaces.

Post performed several other regression analyses. His results generally concluded that Luv and Lab color spaces are not uniform, but they are useful, and that substantial benefits could be

produced by reweighting their axes. These findings are consistent with the other literature discussed in this section.

Further research in this area is still needed. Different tasks, such as reading speeds for characters rather than numerals, and perhaps different response measures need to be investigated.

Perceptual Problems. Walraven (1985a,b) discussed a variety of phenomena that affect the perception of color. His review includes small-field tritanopia, chromatic induction, the Bezold-Brucke effect, the Abney effect, the Helmholtz-Kohlrausch effect, and others. These visual phenomena may or may not be beneficial to human performance on colored displays. While a great deal of research exists which test and describe the conditions of the phenomena, research which relates these various phenomena to real-world tasks is virtually nonexistent. This is another area in which research is necessary if color is to be used appropriately and effectively. Performance for certain color combinations may be due not only to the color contrast (in terms of being able to discriminate the target from the background) but also to the perceptual effects that are occurring because of the colors used, the size of the stimulus, and other parameters.

Cell and Line Failures on Multichromatic Displays

Dot or line cell failures on chromatic displays will result in the loss of one, two, or three of the three primary colors, assuming a three-color primary display system. This is a critical item unique to color displays which obtain color by summation of three primary colors (e.g., R,G,B). If one or two fail, the presented information may not be lost from the display due to the use of nonsaturated colors, but the chromaticity and luminance of the information may change drastically. Therefore, while partial failure is not catastrophic, it may be detrimental to performance. Currently no data exist on this issue.

Areas in Need of Research

The literature review of alphanumeric legibility/readability research has revealed some areas in need of further research. These areas are briefly presented here.

1. While the research dealing with element size, shape, and spacing has provided some insight for optimizing user performance on alphanumeric displays, further investigation is still required before any standards or concrete recommendations can be made. In particular, further comparisons between existing display technologies similar to the work by Snyder and Maddox (1978) are needed.

2. Almost all of the research dealing with alphanumeric legibility/readability has been conducted using upper case characters. The same research questions are valid for lower or mixed case, questions such as optimum matrix size, element size, shape, spacing, angular subtense, font, etc.

3. Viewing angle has been found to affect user performance. LC displays must be viewed within narrow angles, while other light emitting displays may be viewed over wider angles. Cut-off points for optimum performance for the various display technologies need to be established.

4. Few studies have investigated the effects of character orientation. It is feasible that alphanumerics will be rotated from normal (90 deg) on cartographic/symbolic displays; therefore, further investigation of the effects of character orientation on legibility/readability is needed.

5. Snyder (1980) pointed out the need for research investigating the effects of large and small area nonuniformity and edge discontinuity. Since then no research has concentrated on these variables. The effects that nonuniformity may have on various tasks such as reading, recognition, and search should be determined, as well as thresholds. Snyder (1980) listed several variables which should be investigated. For large area nonuniformity, these include viewing angle, mean display luminance, degree of nonuniformity, and shape of the luminance gradient. For small area nonuniformity, the increase or decrease in luminance of individual elements from neighboring elements along with mean display luminance, the number of elements changing in luminance, and the distribution and density of aberrant elements in the display should be investigated.

6. Further investigation into the effects of dot and line failures on matrix-addressed displays is still needed. Abramson and Snyder (1984) provided important data on the effects of line and cell failures for different fonts and upper and mixed case alphanumerics. This research needs to be substantiated further and other variables require investigation, such as matrix size, character size, display polarity, special symbols, and others.

7. The effect that display polarity has on user performance is still unclear. Several studies indicate no difference, while others found substantial differences. How display polarity interacts with other variables, such as contrast, luminance, and dot and line cell failures, requires further investigation.

8. Within the last few years researchers have investigated the quantitative relationship between chrominance and luminance contrast. With the ability to predict perceived color contrast in a uniform color space, further investigations into alphanumeric legibility/readability on color displays and prediction of user performance using the available color metrics is feasible.

Relation to Display Technologies

At the beginning of this report characteristics of the various display technologies were presented. The research on alphanumeric legibility and readability has lead to several recommendations for optimum user performance. It seems appropriate at this time to relate the user performance recommendations to the current display technologies to determine if these recommendations are being followed. Table 18 compares several of the display parameter recommendations discussed previously with each of the display technologies. The recommendation for each variable is located down the rows. Preferred rather than minimum recommendations are presented. For the recommendations listed it appears that the technologies are generally meeting the recommendations at preferred rather than minimum levels.

In many cases it is difficult to make comparisons because whether a display meets the recommendation depends upon the manufacturer; for example, the matrix size used or font can vary depending on who makes the display. Therefore, comparisons for variables such as character size, font, and matrix size are not included; however, it is believed that most of the displays are capable of manipulating these parameters to meet recommendations.

Table 18
Comparison of Display Parameter Recommendations
and Display Technologies

	Contrast	Luminance	Element Size	Element Shape
Recommended Values	(90%)	(35 cd/m ² minimum)	(0.75mm)	(Square)
Display Technology				
CRT	98% at low luminance and low ambient	300	0.07 at 2.35σ	Gaussian
Flat CRT	98%	820	0.35 (est.)	Gaussian
Vacuum Fluorescent	98%	270	0.125	Gaussian
Plasma	95%	600	0.25	variable
EL	92%	3,400	(0.279) ²	selectable
LED	96%	34,000	0.300 X 0.250	round, square
LCD	96%	N/A	0.180 X 0.135	selectable
EC	90%	N/A	(3.175) ²	selectable
EPID	94%	N/A	?	?

Cartographic/Symbolic Research

Cartographic/symbolic research refers to nonalphanumeric information displays created through the use of computer graphics, for example, maps, graphs, or other pictorial information. The distinction is that such displays are neither full alphanumeric displays nor literal images. (Alphanumerics may be on the display as a coding technique.) Computer graphics technology has made it possible to present detailed cartographic and symbolic information on electronic displays. This capability is being used in many military and commercial systems. While there has been a reasonable amount of research investigating the effects of various display parameters on alphanumeric legibility and readability, there is very little research investigating how these variables affect information extraction from cartographic or symbolic displays.

For the most part, design of a pictorial display requires that recommendations based on alphanumeric research be used. Unfortunately it has not been verified that these recommendations can be generalized to nonalphanumeric displays. Researchers must consider that observers performing alphanumeric tasks have the advantage of the redundancy of the English language as well as the familiarity of the alphabet and numbers. This advantage does not exist for tasks which require information extraction from nonalphanumeric symbols. Albert (1975) demonstrated that performance differs for contextual and noncontextual word tasks; therefore, it is very probable that performance using nonalphanumeric displays will differ from that with alphanumeric displays.

The purpose of this section is to discuss the limited research available regarding display parameters and their effects on information extraction from graphic/symbolic displays. Research that has been found in this area deals with symbol resolution and symbol size. No data were found regarding other display parameters, such as luminance, contrast, off-axis viewing, symbol rotation, temporal characteristics, uniformity, or polarity. Therefore, it would be redundant to list each parameter in its own subsection and continuously state that no research was found on that topic. Readers should refer to the Alphanumeric Legibility sections on those parameters, realizing that results may not generalize. Obviously a great deal of research is needed in this area.

Tasks and Dependent Measures

The studies reviewed for this report most commonly used symbolic recognition tasks which require observers to recognize a single symbol on an achromatic or black background. Determining whether observers can recognize symbols or extract information from maps or other complex backgrounds is probably more representative of a real-world task. This does not mean that single symbol recognition data are not relevant; however, results may not generalize to information extraction using a complex background. Contrast, symbol size, and other requirements may be substantially different. Single symbol recognition data also do not take into consideration that observers must typically perform a visual search of the display.

Considering the diversity of tasks that could possibly be performed, the dependent measures commonly used are recognition response time, recognition accuracy, and visual search time. Eye fixations can also be used as a response measure. Visual search response time may be affected by a great many variables, including display density, the number of targets to be searched for, complexity of the background, display noise, color, symbol size, target location, search strategy, contrast, illumination, and many others depending upon the situation. It is beyond the scope of this report to perform a review of all the visual search data. The point is that when using visual search as a task and dependent measure it is important not to confound the variables so that it can be determined which variables are actually affecting performance.

Resolution

Continuous Displays. Semple et al. (1971) stated that CRT resolution requirements are more stringent for symbolic information than for alphanumerics or words. They also stated that resolution must increase as the symbols become more detailed, increasing exponentially with complexity. When symbols are displayed against complex backgrounds, resolution requirements are even more stringent.

Marssetta and Shurtleff (1966) were interested in determining the number of television lines (11, 14, 17, 20, and 23) required to identify 30 military map symbols. The symbols were presented to the subjects in the center of the screen in negative contrast. Results indicated 17 lines with a visual size of 27 minutes of arc was required. For practiced observers a resolution of 11 lines with a visual size of 18 minutes of arc was found to be satisfactory. In this experiment, visual size was confounded with the number of raster lines. It is expected to be that a larger symbol will result in better performance than a smaller symbol. Therefore, it is not possible to tell whether results were

due to the number of raster lines, the symbol size, or an interaction between both. Many technical problems occurred while running the study, which could also have affected the results.

Hemingway and Erickson (1969) studied the relative effects of the number of raster lines and symbol subtense on symbol legibility. Sixteen geometric figures were evaluated. The number of raster lines manipulated were 4.8, 6.3, 7.8, 13.5, 15.5, and 25.6. Symbol angular subtenses were 4.4, 6.0, and 10.2 minutes of arc. The dependent variable was the number of correct responses. The results show that performance improves for all angular subtenses as the number of raster lines increases above 7.8 per symbol height. As the number of symbol lines was decreased, performance was maintained when angular subtense was increased but only for symbols made up of 7.8 lines or more. Comparing their results with other studies, they concluded that performance could be maintained as the number of lines decreased by increasing symbol subtense for subtenses between 7.8 to 16 minutes of arc, after which increasing the subtense did not improve performance. The authors put forth an equation to help determine the number of raster lines and angular subtense necessary for adequate performance:

$$SA = 90, \text{ for } 6 < A < 16, \quad (12)$$

where S is the number of lines per symbol height, and A is the angular subtense in minutes of arc.

Erickson and Main (1966, cited by Hemingway & Erickson, 1969) found that 10 lines per symbol height resulted in 80% accuracy for identifying geometric symbols. Erickson, Main, and Burge (1967, cited by Hemingway & Erickson, 1969) obtained 90% accuracy at 12 lines per symbol height with an angular subtense of 14 minutes of arc.

After a review of the literature, Semple et al. (1971) stated that symbols require 33% to 100% more resolution than alphanumeric characters on the same display for an identification accuracy of 100%. Also, performance increases as the symbol size increases up to 16 minutes of arc.

Matrix Size

Shurtleff (1970c) investigated symbols constructed from matrices of dots to determine matrix size requirements. The height, width, and stroke width of the symbols were held constant. All symbols subtended 22 minutes of arc. Four matrix sizes were investigated: 5 x 7, 5 x 9, 7 x 9, and 7 x 11. Performance was investigated under two levels of symbol overprinting (25% and 50% of symbol height) and no overprinting. Subjects were required to identify each symbol in a 2 x 3 array from left to right. Response speed (correct identifications per minute) for each array and accuracy data were collected. The experiment was conducted in two sessions to determine the effects of practice.

Results indicated that correct identifications per minute improved as matrix size increased from 5 x 7 to 7 x 11. Post-hoc comparisons indicated that for the no-overprinting condition there were significant differences between the 5 x 7 and 7 x 9 matrices and between the 5 x 7 and 7 x 11 matrices. The same results were found for the 25% overprinting condition. For the 50% overprinting condition, only the 5 x 7 and 7 x 9 matrices were significantly different from each other. These results indicate gradual increases in performance as matrix size is increased because no adjacent matrix sizes were significantly different from one another.

For percentage of errors, there were no significant differences among matrices for the no-overprinting condition. For the 25% overprinting condition only the 5 x 7 and 7 x 9 matrices were significantly different from each other. It is rather surprising that there is no difference between the 5 x 7 and the 7 x 11 matrices. For the 50% overprinting condition, significant differences were found between the 5 x 7 and 7 x 9 and the 5 x 7 and 7 x 11.

In general, the results indicate that when there is no symbol degradation performance will increase with increasing matrix size. There was no beneficial effect of increasing matrix size for the percentage of errors response measure. This may have been due to the small sample size used. When symbols are degraded an increase in matrix size is generally required. Also, the analysis indicated that performance improved with practice for both dependent measures under degraded conditions only. The authors recommend using a 7 x 9 matrix size for special symbols.

Cell and Line Failures

No research has been conducted evaluating visual performance under conditions of dot and line cell failure. Differential effects may occur depending upon the type and density of the pictorial display. The effects will probably be quite different than the effects on alphanumeric displays for two reasons. First of all, the display will not have the advantage of familiarity and redundancy found in alphanumeric displays. Secondly, alphanumeric displays have both vertical and horizontal spaces between characters and lines of text which, if affected by dot or line failure, may not cause

interference with performance. With pictorial displays, the spacing involved is not predictable. "On" failures will add noise or perhaps details that may be falsely interpreted as information. The amount and type of failures acceptable on a cartographic or symbolic display need to be investigated.

As mentioned in the Alphanumeric Research section, the effect of dot or line cell failures on multichromatic displays is unknown.

Chrominance

Post (1985) evaluated the effects of color on CRT symbol legibility. Ten symbols were presented individually at three luminance levels (8, 91, and 343 cd/m^2) and 13 chromaticities. Symbols were created from strokes. The symbols' angular subtense was increased from 3 minutes of arc in 0.5 minute steps until the observers could correctly name the symbol, name the hue (either correctly or incorrectly), and until the observers were subjectively "comfortable." All three of these conditions were analyzed as separate response measures.

For the threshold legibility data, the main effects of symbol type, luminance, and chromaticity were significant. The effect of symbols accounted for almost all of the variance. The angular subtenses for this variable ranged from 7.80 to 13.90 minutes of arc. According to Post, the effect of luminance was detrimental. Increasing luminance increased the angular subtense required for detection from 10.24 to 10.96 minutes of arc. The chromaticity effect had a range of only 0.5 minute of arc. Post-hoc analyses indicated no significant differences as a function of the color's purity. (Purity in this experiment referred to a color's distance on a uniform color space rather than excitation purity.) Post believes that another unidentified variable may have covaried with chromaticity to cause the significant results. There was no effect of dominant wavelength.

The comfort legibility threshold measure resulted in significant main effects of symbol and chromaticity as well as an interaction between chromaticity and luminance. The effect of symbol again accounted for most of the variance and the mean subtense values ranged from 15.20 to 21.25 arcminutes. These means are quite a bit larger than those required for detection. Chromaticity post-hoc analyses indicated significant effects of purity but not dominant wavelength; however, the range obtained was only 0.95 minute of arc which, according to Post, is of little practical consequence. The luminance by chromaticity interaction indicated that as luminance increased, chromaticities diminished.

The correct hue-naming measure resulted in significant effects of symbol, luminance, chromaticity, luminance by chromaticity, and luminance by symbol. The chromaticity effect accounted for the most variance. Post-hoc analyses indicated significant effects due to purity but not dominant wavelength. Increasing purity improved hue naming performance; that is, larger subtenses were required to name desaturated colors. In regards to luminance, symbol subtenses increased as luminance increased. The chromaticity by luminance interaction illustrated that threshold size increased for red colors as luminance increased.

The author noted that as luminance increased the stroke width of the symbols increased, which may have been the cause of increased symbol subtense required as luminance increased. Post concluded that "chromaticity has little practical effect on legibility for CRT symbology." It should be noted that this study used color symbols on a black background; therefore, results may not generalize to colored symbols on colored backgrounds. As discussed in the section on Alphanumeric Research, investigations of multichromatic displays are desperately needed.

Areas in Need of Research

It is obvious that a great deal of research is needed to determine recommendations for designing cartographic/symbolic displays for optimum information extraction. A list of all the research possibilities would be endless. Many variables that should be investigated in this area include element size, shape, and spacing; spacing of symbols as well as background information; polarity; symbol sizes and shapes; symbol orientation; dot and line failures; chrominance, luminance, and their contrasts. Many other variables could also be included.

Relation to Display Technologies

While it was possible to compare recommendations for alphanumeric legibility/readability and the display technologies, there has not been any recommendations in the area of cartographic/symbolic research to allow for such a comparison. It should be noted that not all of the display technologies are advanced enough to present cartographic/symbolic information, so that research in this area will have limited near-term application to flat-panel displays.

Literal Image Research

Imaging systems, both analog and digital, are used for many applications, such as photoreconnaissance, space exploration, earth resource management, weather prediction, cartography, archaeology, and medical diagnosis (Avery & Berlin, 1985; Chao, 1983). The technology in this area has advanced from static and dynamic photography to nonphotographic imaging systems which include electro-optical sensors and imaging radar systems. Once the imaging data are collected, they may be stored and presented in a variety of ways. One technique is to encode the image numerically and store and present the image digitally. Digital image processing techniques that enhance or restore the image can then be applied (Avery & Berlin, 1985).

Variables that affect the performance of a human image interpreter can be divided into two categories. The first category includes the knowledge, skills, and abilities of the interpreter. For example, a military photointerpreter may be influenced by the knowledge of additional intelligence information concerning an area that is being reviewed. A second category of variables includes those which actually affect the quality of an image, for example, blur or noise introduced into the image by the imaging system itself.

The purpose of this section is to review the literature dealing with variables that affect image quality and their subsequent effect on information extraction performance. There is not a great deal of literature on this subject. According to Snyder, Turpin, and Maddox (1980) "human factors experiments required to produce quantitative and objective measures of image quality have rarely been conducted in image processing laboratories or in conjunction with image processing programs."

This section of the report will describe four areas of research investigating human performance: analog images, digital images, chrominance, and the effects of dot and line cell failure on literal images.

Analog Imagery

This section describes a series of studies conducted to assess target acquisition performance using static and dynamic film presented on video monitors. These studies relate to interpretation of aerial photography, one of the original tasks for which literal images were used.

Snyder, Keesee, Beamon, and Aschenbach (1974) investigated dynamic target acquisition performance on a constant line rate/video bandwidth system under five different video noise levels. The signal-to-noise levels were 30 (no noise), 20.0, 16.4, 13.0, and 10.4 dB. Noise was obtained by adjusting the noise input to a video mixer. The stimuli were films of simulated flight with a ground speed of 500 ft/s at 10,000 feet altitude. The field of view on the television monitor was 40 degrees vertical by 30 degrees horizontal with a 45 degree boresight depression angle. There was a total of 25 targets which varied in size as well as contrast. Results indicated that as noise level increased, the proportion of correct responses decreased and the proportion of incorrect responses increased. A second analysis was performed on slant range data (slant range to target at the time of a recognition response) and indicated a significant difference between slant range for correct and incorrect responses. The mean slant range was larger for the incorrect responses. The main effect and interactions with noise were not significant. The researchers believe this may have been due to the scoring method.

Snyder (1976) discussed a similar study. In this study target acquisition performance was investigated with five separate video systems, each with different line rate/bandwidth combinations under five levels of noise. Table 19 lists each of the video system line rates/bandwidths and the five noise levels used for each.

Subjects were required to search for three targets under each noise level on three different missions. The missions varied by flight geometry. The depression angles and velocities of the three missions were 45 degrees, 500 ft/s; 23 degrees, 500 ft/s; and 23 degrees, 3000 ft/s. Percent correct responses and ground range to target for correct responses were analyzed separately for each of the five video systems. For all five systems the main effect of noise and mission were significant for the percent of correct responses. In general, as noise increased, the percent of correct responses decreased. Performance was superior for the 45 degree, 500 ft/s mission condition for all five video systems. As the depression angle decreased or as velocity increased, correct responses decreased.

Analyses were also performed for each of the five systems on the target's ground range at the time of a response, either correct, incorrect, or no response. However, 12 separate analyses were performed and readers are referred to the study for specific results. In general, the results indicated that the 23 degree, 500 ft/s mission produced the largest ground ranges while the 45 degree, 500 ft/s produced the shortest.

Table 19

**Video System Linerates, Bandwidths, and
Noise Levels Used by Snyder (1976).**

Video System Linerate/Bandwidth (M Hz)	Noise Levels (in MV)				
525/8	0	10	20	50	100
525/16	0	7	14	35	70
945/16	0	7	14	35	70
1225/8	0	5	10	25	50
1225/32	0	5	10	25	50

All five systems could not be compared in one analysis because the noise levels for each were not equated; however, a comparison could be made at the zero noise level. An analysis of variance was performed for the number of correct responses at the zero noise level and the main effects of mission and video system were significant. Again, the 45 degree, 500 ft/s mission was superior to the other mission conditions. Post-hoc comparisons indicated significant differences between all pairs of mission means. Post-hoc comparisons also indicated a significant difference between the 525/16 and the 945/16 systems. No other differences were significant. It appears that increasing line rate significantly increases performance. However, the authors state that "there is a diminishing benefit to increasing line rate much over 1000 lines" and that this finding is in agreement with other studies. The differences among video systems at the zero noise condition for the ground range performance measure were not significant.

In summary, results of correct responses and ground range indicate a trade-off between depression angles. For large depression angles, targets are larger and therefore easier to detect under all noise levels; however, the ground range (or acquisition range) is reduced using larger angles. Increasing velocity also decreases the ground range. Also, increasing video noise level decreased the number of correct responses from all display systems.

Gutmann, Snyder, Farley, and Evans (1979) conducted two studies (dynamic and static imagery) which investigated target acquisition performance. These studies were very similar to those previously discussed. They found that noise levels did not affect correct responses for the static display experiment, but that noise did affect performance in the dynamic experiment. Increases in noise led to decreases in target acquisition. They also found that as target size increased correct responses increased and the targets were acquired at greater ranges. Target sizes were defined as small, medium, and large; therefore, it is not possible to make target or symbol size recommendations from this report.

Results of these various studies indicate that increases in video noise and decreases in target size result in decreased target acquisition. Snyder (1974) discussed two reasons for this effect. First, video noise masks smaller targets and target details. Noise in the spatial frequency range of the target and below that of the target is more detrimental than is noise in frequencies above the target. There is more noise below the smaller target's higher spatial frequencies (Keesee, 1976, cited by Snyder, 1976). Also, it is possible that visual search strategies changed as noise increased. Snyder and Taylor (1976) found that as clutter increases subjects use shorter visual interfixation distances, which increase search times. Increasing target size and as well as depression angle (or equivalently, symbol size) will enhance correct response performance in noisy and zero noise video displays.

Digital Imagery

The Human Factors Laboratory at Virginia Polytechnic Institute and State University conducted a 5-year research program evaluating the quality of digitally derived imagery used in military photointerpretation operations. Specifically the effects of blur and noise were investigated.

This research effort developed a large digital database, established subjective and objective measures of image quality, compared hard-copy and soft-copy displays, evaluated different enhancement techniques, and compared different image quality metrics (discussed in Section 4). In all the studies conducted, both subjective and objective measures of image quality were taken. The subjective measure was a modified 10-point NATO scale (see Snyder, 1983). The objective measure required photointerpreters (PI) to answer a series of specific questions about essential elements of information (EIs) found in the images. The task was similar to real tasks performed by the photointerpreters.

The soft-copy experiments were conducted on two high resolution monochrome CRTs. Three levels of blur (26, 94, and 364 micrometers) and five levels of noise (SNRs of 208, 100, 50, 25, and 12.5) were investigated. An ANOVA was performed on the objective information extraction measure. Significant effects of both blur and noise were found. As blur increased the percent correct EIs decreased almost linearly. Percent correct EIs also decreased as noise was increased. An interaction between noise and blur indicated that the effect of noise was reduced at the largest blur level, but it was not a statistically significant result (Snyder, 1983).

Results of the subjective NATO scale found that increases in blur caused decreases in the NATO scale values, as did increases in noise. An interaction between blur and noise revealed that the effect of noise was reduced at larger blur levels. Objective and subjective measures were found to correlate, $r = 0.965$, $p = 0.0001$. This finding indicates that subjective measures may be used accurately to predict performance. Researchers found very similar results for the hard-copy experiments (Snyder, Turpin and Maddox, 1980; Snyder, Shedivy, and Maddox, 1980).

Chao (1983) investigated the effect of 10 different image enhancement restoration techniques (listed in Table 20) on blurred and noisy images using the subjective and information extraction performance measures.

Table 20

Image Enhancement/Restoration Techniques
Used by Chao (1983)

-
- Contrast modification
 - linear stretch
 - adaptive contrast stretch + noise filter
 - Deblurring
 - unsharp masking + noise filter + linear stretch
 - Laplacian filter + noise filter + linear stretch
 - Noise removal
 - noise filter
 - neighborhood averaging + linear stretch
 - adaptive noise filter + linear stretch
 - Deblurring and noise removal
 - Wiener filter + noise filter + linear stretch
 - Miscellaneous operations
 - noise filter + linear stretch
 - no processing
-

Due to various time and resource constraints, the information extraction task required that the blur and noise variables be combined to form one independent variable (see Figure 18), and only five of the restoration/enhancement techniques were used. Subjects (10), scenes, blur/noise, and processing techniques were combined into two 5 x 5 Greco-Latin square designs. An ANOVA was performed and the main effects of scene and blur/noise were significant. Post-hoc results indicated that of the 10 blur/noise levels only three (364/12.5, 26/12.5, and 26/50) were different from other levels and these blur/noise levels yielded the lowest EEI scores. These results indicate that high blur/high noise images degraded performance the most.

		BLUR (μM)				
		26	50	94	184	364
SNR (NOISE)	200	1		4		7
	100					
	50	2		5		8
	25					
	12.5	3		6		10

Figure 18. Combined blur and SNR (noise) used by Chao (1983).

The subjective scaling study investigated all 10 enhancement/restoration processes, three blur levels, three signal to noise ratios, and five scenes. Ten PIs scaled the scenes for all levels. A $3 \times 3 \times 5 \times 10 \times 10$ ANOVA was performed. All main effects and interactions except two were significant. Due to the complexity of the results, a general summary of the findings will be presented. Readers are referred to Chao (1983) for further information. Chao stated that increasing the degradation of the images by either blur or noise consistently reduced the judged interpretability. There were no significant differences in perceived interpretability between the two lowest levels of blur (26 and 94 μm), or between the two lowest levels of noise (50 and 200 SNR). She concluded that this may have been due to closely spaced degradation levels.

In regards to the enhancement/restoration techniques deblurring techniques, generally helped with blur removal and noise reduction techniques generally helped noise degraded images, as would be expected. Table 21 (Chao, 1983) summarizes the relative effectiveness of each technique at improving degraded images.

In summary, the research illustrates that noisy and blurred digital images affect user performance and that various techniques are available to aid in enhancing and restoring images to aid interpreters. It must be noted, however, that these data were obtained from CRT, not flat-panel, displays, although the images were digitally stored and processed. Images on matrix-addressed flat-panel displays and any differential effects due to the display technology have not been investigated.

Table 21
 Usefulness of Restoration/Enhancement Processes for
 Improving Images Degraded by Blur and Noise (Chao, 1983).

Noise	Blur		
	Low	Intermediate	High
Low	unsharp masking noise filter neighborhood average linear stretch Wiener filter	unsharp masking Laplacian filter noise filter linear stretch Wiener filter	Laplacian filter
Intermediate	noise filter neighborhood average noise filter + linear stretch Wiener filter	unsharp masking Laplacian filter noise filter adaptive noise filter noise filter + linear stretch Wiener filter	unsharp masking Laplacian filter noise filter noise filter + linear stretch Wiener filter
High	noise filter neighborhood average adaptive noise filter adaptive contrast stretch Wiener filter	unsharp masking noise filter neighborhood average adapt. noise filter adaptive contrast stretch noise filter + linear stretch Wiener filter	unsharp masking Laplacian filter noise filter neighborhood average adaptive noise filter adaptive contrast stretch noise filter + linear stretch Wiener filter

Cell and Line Failures

There has been no published research to date which investigates dot or line failure on literal image displays and their effect on human performance. The same considerations discussed in the cartographic/symbolic section apply to literal image research. Also, with literal images it is likely that every element on the display will be utilized to create an image; therefore, "off" failures may be more detrimental than "on" failures. The amount and type of failures acceptable on literal image displays need to be determined.

The effects of dot and line failures on multichromatic literal images may be substantially different than the effects on monochromatic displays. As pointed out in the earlier sections, differential failure of one or two of the three primary colors may cause chromaticity and/or luminance changes of the information.

Chrominance

The studies reviewed in this section were all achromatic, varying in gray scale levels. Studies which investigate the effects of blur and noise or contrast on colored images have not been found. Color is used as an enhancement technique and no data were found which address whether enhancing images with color results in improved image interpretation. Research is required to determine the effect of color on image interpretability.

Areas in Need of Research

The research discussed in this section has primarily dealt with the effects of image degradation (specifically blur and noise) of literal images on information extraction tasks. Other areas in need of research are listed below.

1. The studies reported in this report were all performed on gray scale images. Often color is used as an enhancement/restoration technique. It is often assumed that color will aid in target acquisition because of increased contrast; however, empirical research has not been conducted. Information extraction performance on multichromatic displays should be investigated.

2. Element size, shape, and spacing for literal image displays should be investigated to determine optimum sizes.

3. The effect of dot and line failures on information extraction for literal image displays requires investigation. It is possible that literal images will be affected differently than alphanumeric or cartographic/symbolic displays.

Relation to Display Technologies

As with cartographic/symbolic displays, concrete recommendations for various display parameters were not found; therefore, comparisons cannot be made until further research is conducted.

SECTION 4

IMAGE QUALITY METRICS

As described above, many research studies have been carried out in an attempt to predict the effects of various display parameters upon the information transfer from soft and hard copy displays. In addition to this, dozens of studies have been performed to predict the effects of the same and other display parameters upon display image quality. These studies have ranged from subjective, or perceived image quality, to objective forms of visual performance such as object recognition time and accuracy of response. Many investigations have evaluated the effect of only one display parameter, while a few have attempted to investigate the interaction effects of several display parameters.

As pointed out by Snyder (1973, 1980), empirical research which would investigate all the possible and important design variable interactions is an impossibility. Because of this, a research strategy that mixes both mathematical modeling approaches and experimentally derived data needs to be adopted to arrive at a useful prediction of the effects of the numerous combinations of display design variables. Unitary image quality metrics have been developed to account for most of the important display design variables which influence subjective image quality or observer information extraction performance, or both.

Snyder (1980) has divided the development of image quality metrics into spatially continuous and spatially discrete forms. Spatially continuous displays are displays which have continuous sampling in both dimensions. That is, the displays are not broken by noninformation bearing borders or edges. Examples of these displays are photographic images or nonraster CRTs.

Spatially discrete displays have artificial lines or edges between the information-bearing image elements. All dot-matrix displays having separate XY cells fall into this category. These displays and their associated image quality metrics are the major topics of this review. There are also hybrid displays which have one continuous dimension and one discrete image dimension. The monochrome television display is a good example of this kind of display, with continuous information horizontally along the raster line and discrete information vertically.

Many of the image quality metrics to be discussed in this review have been developed for and tested on continuous image and hybrid image displays and do not apply directly to the evaluation of flat-panel discrete displays. However, these measures form the basis of nearly all quality metric concepts and are important to the discussion. As will be shown, many of the image quality metrics for continuous displays have been adapted to the evaluation of discrete image displays.

Before discussing the applicable image quality metrics in detail, a few general comments common to the discussion of image quality should be reviewed. Snyder (1985) has described the term "image quality" as being used in two general contexts: (1) that dealing with physical measures of the image itself and with little or no regard for the ability of the observer to obtain information from the image; and (2) that dealing with perceived or measured quality from the human observer, sometimes with little regard for the physical characteristics of the image.

The two physical measures of image quality which have been used the most are based on either the modulation transfer function (MTF) or some bivariate error statistical (pixel error) measure. MTF-based measures determine the displayed contrast in an image as the function of the size of objects in the image. Pixel error measures relate the intensity distributions of an image to assumed ideal intensity distributions or relate an original image to a degraded version of the image such that the differences in the statistical intensity distributions are a measure of the degradation of the image quality. Taken by themselves, these measures of image quality physically describe the image in terms of either measured or calculated luminance units. No regard is given to how these measures relate to an observer's ability to gather information from the displayed image.

In contrast to the pure physical measures of image quality, behaviorally validated measures emphasize the visual performance or perception of the observer. This performance is then related empirically to physical characteristics of the image. Such validated measures of image quality can be used to develop models which can predict an observer's performance in terms of information extraction from displays. Because of this, it is both meaningful and useful to conduct and apply research that relates physical measures of image quality to observer performance. One of the primary goals of the current research program is the development of such models with the specific application to matrix-addressed displays and their associated failure modes. Throughout the review

of image quality metrics, emphasis is placed on those metrics which have been behaviorally validated and show promise for use as display quality metrics for matrix-addressed displays.

The review of image quality metrics is organized into three sections. First, the existing display image quality metrics pertinent to the design variables and failure modes of flat-panel displays will be discussed. The appropriate formulae, original references, and limitations of the selected metrics will be given.

Second, the image quality metrics described in the first section which have been behaviorally validated will be reviewed thoroughly. Each metric will be evaluated in terms of type of display used (CRT, matrix-addressed, photographic image, etc.), performance measure (response time, recognition latency, etc.), and type of information displayed (alphanumeric, literal image, or cartographic/graphics).

Third, the areas in need of research will be pointed out. These will be limited to those areas which fulfill the purposes of this research program.

MTF-Based Metrics

The first set of image quality metrics to be considered are derived from the Modulation Transfer Function (MTF). The MTF is based upon the theory of linear systems analysis and the mathematics of Fourier transformations. The concept of linear systems analysis permits one to determine the extent to which any component or system of components can transmit a signal. In the transmission process, some of the signal's amplitude is often lost, due to limitations of the transmission system, and this loss is measurable if the measurement is made under the proper circumstances. The MTF is a way to measure this degree of fidelity of transmission in a display device.

As stated above, it is an impossibility to perform the virtually infinite number of experiments which would be necessary to describe a specific display's capability for reproducing objects of varying shapes, varying sizes, varying contrasts, and under the conditions of varying adapting luminance levels which represent the many different uses of displays. The very powerful techniques of Fourier analysis can be used to much abbreviate the requirement for such a multitude of experiments. Specifically, Fourier analysis states that any repetitive waveform can be analyzed into a number of component frequencies, with each component frequency having a specific amplitude and phase relationship. If all the frequencies are appropriately combined with their respective amplitudes, the resulting summation is the original repetitive waveform, however complex it may have been (Snyder, 1980).

Of considerable importance in the design of any display system is the fact that the high spatial frequency information must be preserved if the high frequency information is critical to the performance of the task by the observer. Thus, a Fourier analysis of the displayed information can be used to determine if the necessary high frequencies are present, and at what amplitude they are represented. If their amplitudes exceed the observer's threshold, then the information is detectable and potentially useful.

Typically, as the frequency of some input to a display system increases, the amplitude of the resulting image will tend to be reduced. The amplitude of the displayed information can be plotted relative to the amplitude of the input information to determine the degree to which a given imaging system can transfer the spatial frequencies contained in the input signal to the image plane of the display. When the relationship is expressed in the form of modulation and the transfer function is described in ratio form then one has the basis for the technique known as modulation transfer function analysis (Snyder, 1980).

The displayed modulation (M) is the ratio of the difference, peak to peak, of some sinusoidal signal as displayed, to the sum of the maximum and the minimum of that signal. This relationship is shown in Equation 13.

$$M(\omega) = \frac{L(x_1)_{\max} - L(x_2)_{\min}}{L(x_1)_{\max} + L(x_2)_{\min}}, \quad (13)$$

where ω refers to the spatial frequency of the measured sine-wave pattern and $L(x_1)$ and $L(x_2)$ denote the intensity or luminance of the sine-wave pattern at display coordinates x_1 and x_2 respectively.

The modulation transfer factor is the ratio of the modulation out of the system to the modulation into the system. Equation 14 shows this relationship.

$$T(\omega) = M_o(\omega)/M_i(\omega), \quad (14)$$

where $T(\omega)$ is the modulation transfer factor at spatial frequency ω and $M_o(\omega)$ and $M_i(\omega)$ are the respective output and input modulations. When the display is unable to pass a spatial frequency without attenuating it, $M_o < M_i$ and the modulation transfer factor is less than unity.

Connecting the modulation transfer factor values for all spatial frequencies forms a continuous function, the modulation transfer function as shown in Figure 19. For an in-depth discussion of the MTF in optical systems, see Gaskill (1978).

For a discrete (digital) image the Fourier transform of a $N \times M$ digital image is given by

$$F(\omega, \nu) = \sum_{x=0}^{n-1} \sum_{y=0}^{m-1} L(x,y) \exp\{-j2\pi\omega x + \nu y\}, \quad (15)$$

where $L(x,y)$ is the image intensity at spatial location x, y in rectangular coordinates, $F(\omega, \nu)$ is the Fourier transform coefficient at spatial frequency (ω, ν) , and N and M are the numbers of discrete image samples along the x and y axes. In subsequent formulae the limits of summation will be omitted unless they differ from those in equation 15.

Generally, $F(\omega, \nu)$ is a complex function composed of a real part and an imaginary part indicated as

$$F(\omega, \nu) = \text{Re}(\omega, \nu) + j\text{Im}(\omega, \nu), \quad (16)$$

where $j = (-1)^{0.5}$. The amplitude of each sine-wave component is given by

$$A(\omega, \nu) = |F(\omega, \nu)| = \{\text{Re}(\omega, \nu)^2 + \text{Im}(\omega, \nu)^2\}^{0.5}, \quad (17)$$

and the phase angle is given by

$$P(\omega, \nu) = \tan^{-1}\{\text{Im}(\omega, \nu)/\text{Re}(\omega, \nu)\}. \quad (18)$$

The two-dimensional modulation spectrum (MTF) of a digital image is computed from the normalized Fourier amplitude spectrum,

$$M(\omega, \nu) = A(\omega, \nu)/A(0,0). \quad (19)$$

The two-dimensional MTF of an imaging system as expressed in terms of the Fourier coefficients is given as

$$T(\omega, \nu) = \frac{M_o(\omega, \nu)}{M_i(\omega, \nu)} = \frac{A_o(\omega, \nu)/A_o(0,0)}{A_i(\omega, \nu)/A_i(0,0)}. \quad (20)$$

It may be helpful to elaborate further on the MTF and its relationship to other display concepts and measurements. The MTF is the normalized Fourier transform of the line spread function often used in photographic image analysis for analog images (Dainty & Shaw, 1974; Gaskill, 1978). The line spread function is the spread of an image (output) of an infinitely narrow line input. When the image of the narrow line is formed, the measured image is no longer a sharp line but has "rounded" edges-- the intensity profile is spread or blurred by the imaging device. The line spread function defines the profile of the resulting image and can be obtained by directly measuring the luminance distribution. Alternatively, the luminance or intensity distribution can be measured across a displayed "knife" edge and differentiated. The differentiated edge is the line spread function.

The line spread function and the normalized MTF are the inverse of one another. In addition, the width of the line spread function is inversely proportional to the passband of the MTF. Thus, either concept may be used to characterize the physical performance of an imaging device or component in one dimension. This mathematical similarity is used as the basis of a number of proposed metrics of image quality.

The composite MTF of an imaging system can be determined simply from the cascading of the MTFs associated with n components of the system. That is,

$$T_s(\omega, \nu) = \prod_{i=1}^n T_i(\omega, \nu), \quad (21)$$

where $T_s(\omega, \nu)$ is the system modulation transfer factor at spatial frequency (ω, ν) and each $T_i(\omega, \nu)$ is the modulation transfer factor for a component of the system.

The mathematical definitions given above describe the MTF concept for a digital image rather than an analog one. This is deliberate in that all flat-panel displays of interest in this review are

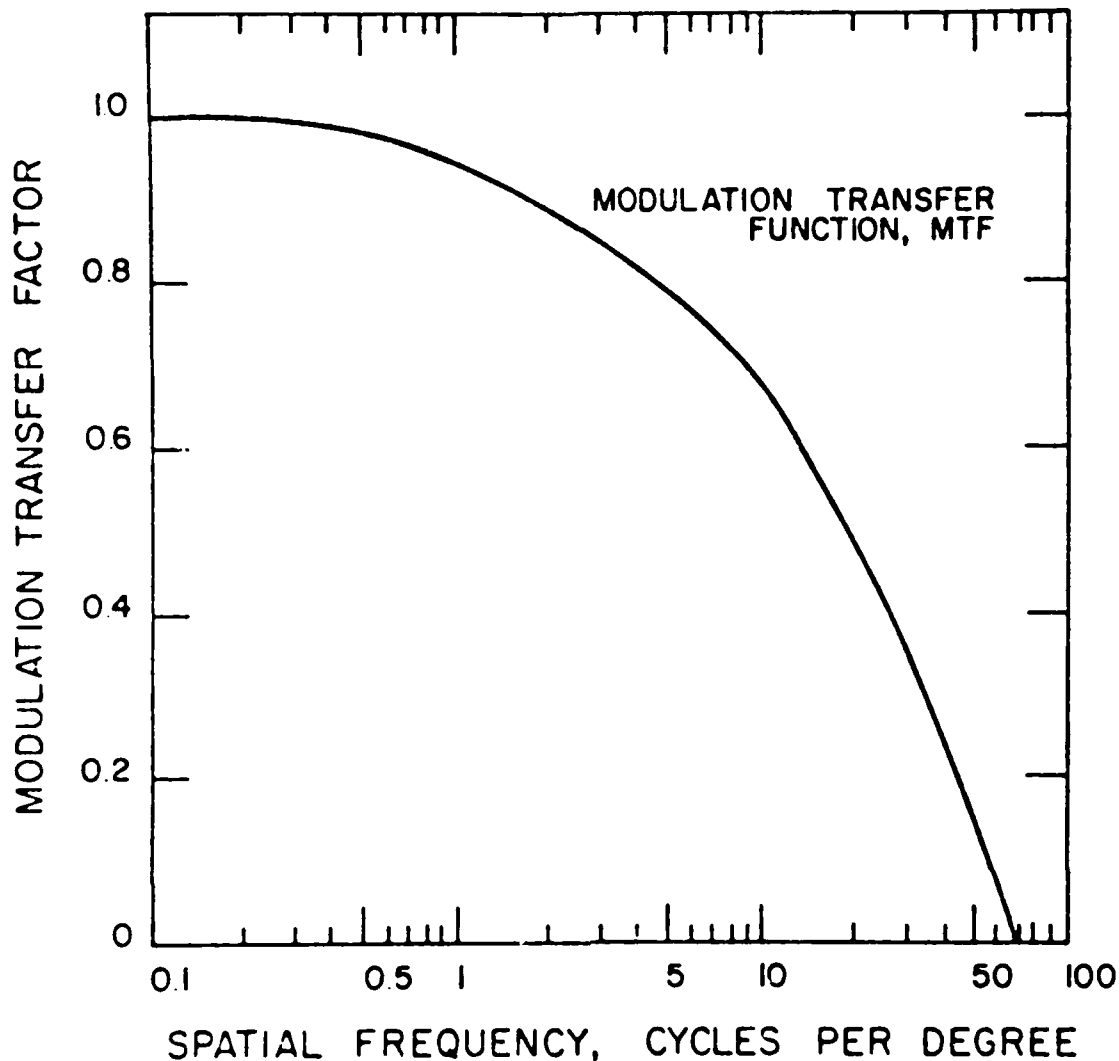


Figure 19. Modulation transfer function.

digitally addressed and are composed of discrete pixels rather than of continuous image information. The exception to this generalization is the CRT image, particularly the raster-scanned image in which the along-raster dimension is a continuous image and the across-raster image is discrete. This distinction is not critical to an appreciation of the MTF-based measures of image quality. The distinction is important only in the calculation of the MTF. For more information and discussion of the calculational differences see Gaskill (1978).

Equivalent Passband (EP)

Schade (1953) developed the concept of EP as the means to describe the quality of a television signal. This metric expresses the width of a system MTF. EP is the equivalent bandwidth of a rectangular MTF which contains the same total sine-wave power as does the actual MTF of a system. That is, it is the cut-off frequency of the perfect filter passing the same power. This concept is illustrated in Figure 20. The EP metric is defined mathematically as

$$EP = \Delta\omega\Delta\nu \sum \sum \{T_s(\omega, \nu)\}^2. \quad (22)$$

where $\Delta\omega$ and $\Delta\nu$ denote the frequency spacing used in numerical integration.

EP is a measure of the "sharpness" in an image. Although this factor certainly relates importantly to the perceived quality of an image, this metric is limited in that it does not take into account the "error" data in an image caused by correlated or uncorrelated noise from any source (Snyder, Keese, Beamon, & Aschenbach, 1974). Nor does it take into account other display/observer system parameters which have been determined to affect observer performance but not the value of EP, such as the angular size of the display (Task, 1979).

For purposes of interpretation, larger values of the EP metric indicate greater imaging capacity and greater system quality.

Equivalent Width (EW)

Bracewell (1965) defined another metric based upon the notion of width as the area under a function divided by its central ordinate. Bracewell demonstrated that the equivalent width (EW) of a function is equal to the reciprocal of the equivalent width of its transform. It follows that the equivalent width of the line spread function is given by the reciprocal of the width of the system MTF (Beaton, 1984), as

$$EW = \frac{T_s(0,0)}{\Delta\omega\Delta\nu \sum \sum T_s(\omega, \nu)}, \quad (23)$$

where $T_s(0,0) = 1$ for normalized system MTFs. Small values of EW indicate greater system quality than do larger values.

Squared Spatial Frequency (SSF)

While the Equivalent Passband concept described above has its limitations, it has influenced the thinking of many image quality researchers in the development of measures based on weighted MTFs or integrated, weighted MTFs. Hufnagel (1965) suggested a weighting scheme which uses the squared spatial frequency (SSF) argument of the system MTF. This is given by

$$SSF = \Delta\omega\Delta\nu \sum \sum T_s(\omega, \nu)(\omega^2 + \nu^2). \quad (24)$$

In evaluating system quality, larger values of SSF indicate a slower approach to a modulation transfer factor of zero, which implies greater system bandwidth and therefore quality (Beaton, 1984).

Strehl Intensity Ratio (IR)

Another metric based on weighted MTFs is the Strehl Intensity Ratio. It is the ratio of the maximum spread function values for an imaging system to that of an equivalent aberration-free system (Linfoot, 1960). It is defined as

$$IR = \frac{\Delta\omega\Delta\nu \sum \sum T_s(\omega, \nu)}{\Delta\omega\Delta\nu \sum \sum T_i(\omega, \nu)}, \quad (25)$$

where $T_i(\omega, \nu)$ denotes the MTF of the ideal system. The Intensity Ratio is no more useful than is the EP concept in evaluating the quality of images containing noise.

Perceptually Weighted System Metrics

The metrics discussed so far define the quality of a display device strictly in terms of the system MTF. In doing so these metrics define the resolution of a display in quantitative terms. The resolution of a display device is only one of the parameters important to human judgements of quality. Other metrics have been developed to take into account the capabilities of the human visual system and relate them to the objectively defined system resolution.

Thus far in this discussion the MTF concept has been applied to displays, and not to the visual system. However, if the visual system is considered to have an input, a spatially varying sinusoidal grating, and an output, the perception of that sinusoid, then the notion of the MTF of the visual system can be explained in terms of that compatible with display devices.

Specifically, linear systems analysis, applied to the visual system, permits us to analyze any displayed pattern into its component frequencies, amplitudes, and phase relationships. It assumes that the visual system behaves as a Fourier analyzer, decomposing complex patterns into

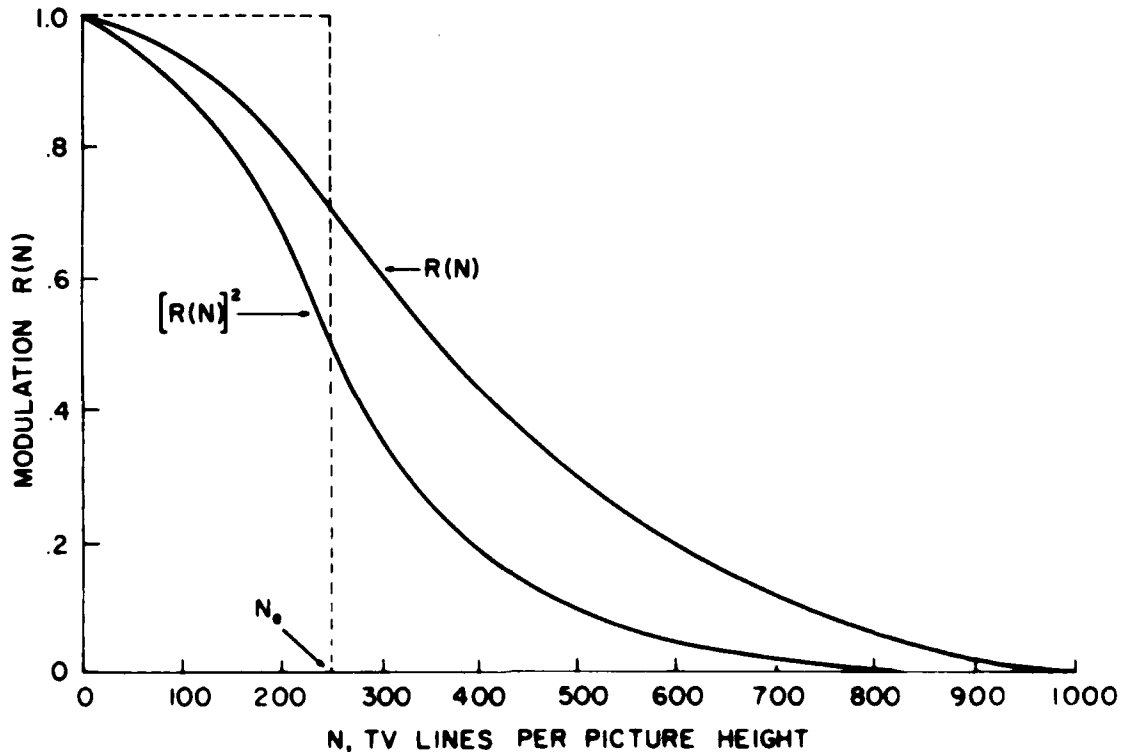


Figure 20. Noise equivalent passband.

frequency/amplitude/phase combinations and responding to each of the component frequencies independently. Given that a complex display can be Fourier analyzed and given also that the visual system behaves as a Fourier analyzer, then it is a tractable analytical matter to determine which of the frequencies contained in the display are "visible" (i. e., above the visual threshold).

Numerous experiments (e.g., Campbell & Robson, 1968) have carefully demonstrated that the visual system, in fact, behaves as a Fourier analyzer in the spatial domain, at least to an adequate first approximation. As a result, we can indicate the sensitivity of the visual system to a standard pattern which is used in linear systems analysis and then compare this sensitivity to the frequency spectrum of the displayed information to determine the sensitivity of the visual system to that information. The standard pattern used for this purpose is the sine-wave grating. In visual threshold experiments using this approach, the sine-wave grating is varied in spatial frequency (cycles per unit display distance or, more usefully, cycles per visual degree of angular subtense). The observer adjusts the modulation of the grating to a threshold criterion. Assuming that the displayed modulation is uniform, the modulation needed to reach a threshold response is then an indication of the sensitivity of the observer to that spatial frequency.

When plotted as threshold contrast as a function of spatial frequency, the resulting function is termed the contrast threshold function, or CTF. The typical CTF has a minimum in the region of 3-5 cyc/deg, with increasing modulation required to reach a threshold response at both higher and lower spatial frequencies. Figure 21 illustrates a typical CTF for normal, healthy, corrected adult eyes. Also illustrated is the estimated deviation from this typical curve for 90% of the population.

The CTF has become a basis for the quantitative analysis of display quality, as will be shown in the discussion of the remaining metrics. Some metrics also use the contrast sensitivity function (CSF), which is the inverse of the CTF. It is important to realize, however, that the CTF is altered by various display surround conditions pertinent to display operational design and usage. For a thorough discussion of the variations in the CTF due to these parameters see Snyder (1980).

The metrics which have been described so far can be expanded to include the CSF as follows:

$$PEP = \Delta\omega\Delta\nu \sum \sum T_s(\omega, \nu)^2 C_i(\omega, \nu)^2. \quad (26)$$

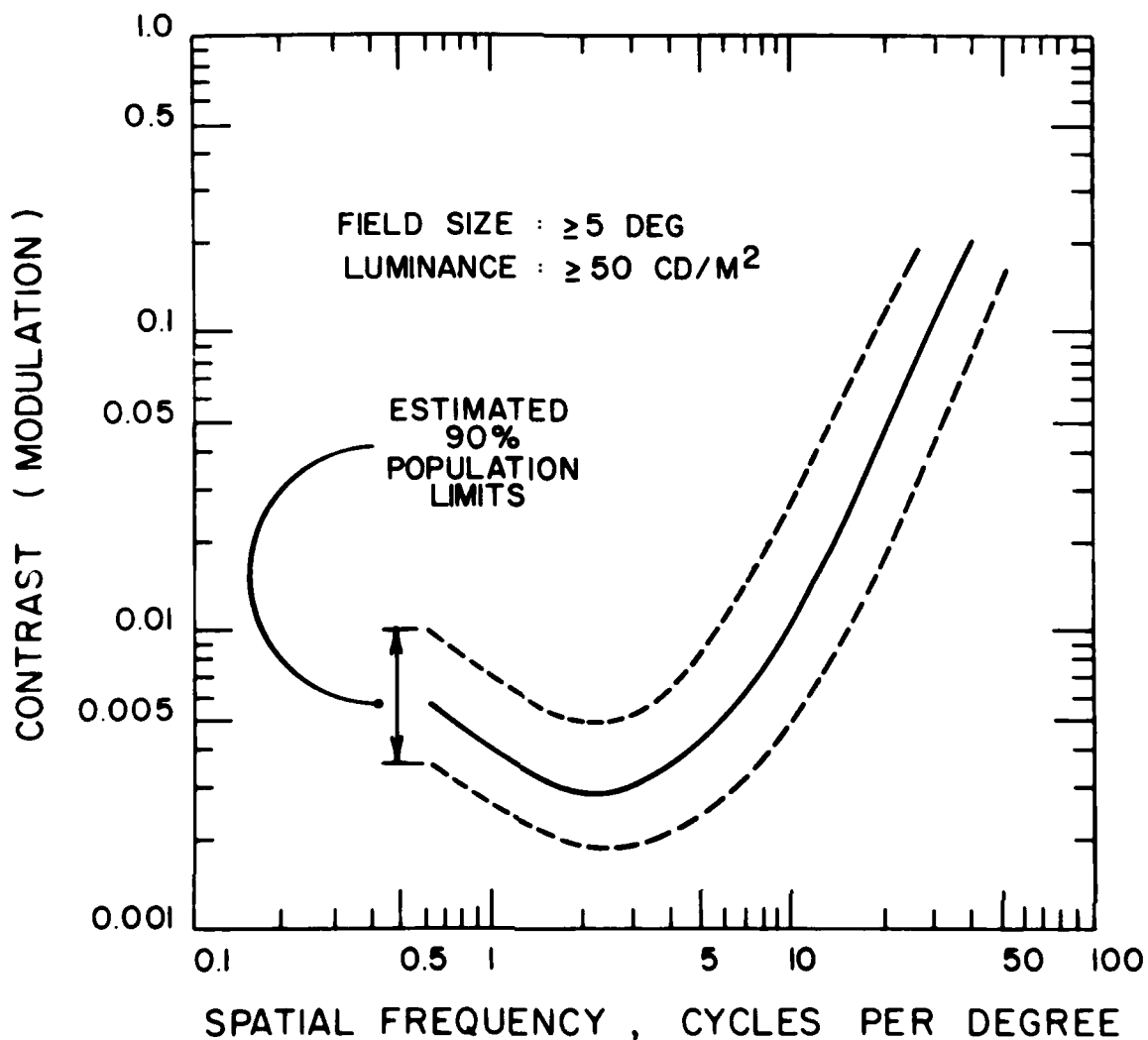


Figure 21. Visual contrast sensitivity function.

$$PEW = T_s(0,0)C_i(0,0)/\Delta\omega\Delta\nu \sum \sum T_s(\omega, \nu)C_i(\omega, \nu), \quad (27)$$

$$PSSF = \Delta\omega\Delta\nu \sum \sum T_s(\omega, \nu)C_i(\omega, \nu)(\omega^2 + \nu^2), \text{ and} \quad (28)$$

$$PIR = \frac{\Delta\omega\Delta\nu \sum \sum T_s(\omega, \nu)C_i(\omega, \nu)}{\Delta\omega\Delta\nu \sum \sum T_i(\omega, \nu)C_i(\omega, \nu)}, \quad (29)$$

where $C_i(\omega, \nu)$ denotes the visual CSF determined under the i th viewing condition. The prefix "perceptual" (P) has been added to indicate the perceptually weighted form of the metric.

Modulation Transfer Function Area (MTFA). One of the most researched metrics of image quality takes into account the MTF of the imaging system or display as well as the CTF of the visual system. This concept was originally suggested by Charman and Olin (1965) and has since been

evaluated by several researchers (e.g., Blumenthal and Campana, 1981; Borough, Fallis, Warnock, & Britt, 1967; Snyder, 1973, 1974, 1976; Snyder et al., 1974; Task, 1979).

The MTFA is illustrated in Figure 22, and is defined in the one-dimensional case as the area between the MTF and the CTF, between zero spatial frequency and the crossover frequency of the two curves. It is often conceptualized as a "signal minus noise" integrated over all usable spatial frequencies. Furthermore, the crossover spatial frequency is the "limiting resolution" of the imaging device.

Mathematically, the MTFA is defined as

$$\text{MTFA} = \Delta\omega\Delta\nu \sum_{\omega=-f}^{\omega=f} \sum_{\nu=-f}^{\nu=f} \{T_s(\omega, \nu) - T_c(\omega, \nu)\}, \quad (30)$$

in which $T_c(\omega, \nu) = C_i(\omega, \nu)^{-1}$. The spatial frequency f = the limit of summation where the system MTF has the same value as the CTF.

The rationale behind the MTFA is simple. It summarizes the excess signal (MTF) over the threshold requirement (CTF) of the visual system over all usable spatial frequencies. It further assumes that the area is homogeneous in image quality; that is, that the excess of MTF over the CTF is uniformly important or isotropic for all spatial frequencies and for all amounts of modulation above the threshold requirement. This assumption has been questioned and tested experimentally but with no substantial and consistent improvement in the concept.

As originally proposed by Charman and Olin and as used by subsequent researchers, the MTF is measured for a given system in the traditional fashion. The CTF is determined either experimentally or analytically. The CTF is used to account for differences in viewing conditions, the gamma of the display or imaging system, and the noise content of the display. In general, as gamma increases or as noise decreases, the CTF is lowered to provide a larger MTFA value. For the rationale and quantitative approach to these manipulations, see Snyder (1973; 1980).

Gutmann, Snyder, Farley, and Evans (1979) tested the isotropic assumption of the MTFA and found that the assumption was unsupported for systems having atypical MTFs. For systems having similarly shaped MTFs, the correlations between MTFA and observer performance are typically quite high. Beamon and Snyder (1975) have suggested that the area immediately above the CTF is of greater importance to the observer than the area well above the CTF. Stated differently, it is critical to have adequate signal (modulation) above that minimally required for detection (CTF), but additional increases in this excess of MTF over CTF are of less value in most real-world tasks. The next quality measure is an attempt to overcome this problem in the MTFA.

Gray Shade Frequency Product (GSFP). Task and Verona (1976) proposed a nonlinear transform of the MTFA to weight the area near the CTF more heavily than the area well above the CTF in an attempt to produce a perceptually isotropic measure of system quality. This transformation uses as its logical basis the assumption that the visual system can be modeled as a logarithmic amplifier which sees modulation proportional to the logarithm of the modulation. They transformed the modulation axis into "just-noticeable differences" or "shades of gray" (G), by the formula

$$G = 1 + \frac{\log_{10}\{(1 + M)/(1 - M)\}}{\log_{10}\{2.0^{0.5}\}}, \quad (31)$$

where the numerator is the modulation and the denominator is the approximation of the modulation difference between successive shades of gray. However, the denominator term does not represent a true JND of luminance.

The two-dimensional GSFP is defined as

$$\text{GSFP} = \Delta\omega\Delta\nu \sum_{\omega=-f}^{\omega=f} \sum_{\nu=-f}^{\nu=f} G\{T_s(\omega, \nu) - T_c(\omega, \nu)\}. \quad (32)$$

Integrated Contrast Sensitivity (ICS). van Meeteren (1973) proposed another approach to "perceptually weighting" the system or display MTF and then cascading it with the visual system CTF. In this approach ICS is defined as

$$\text{ICS} = \Delta\omega\Delta\nu \sum \sum T_s(\omega, \nu)C_i(\omega, \nu). \quad (33)$$

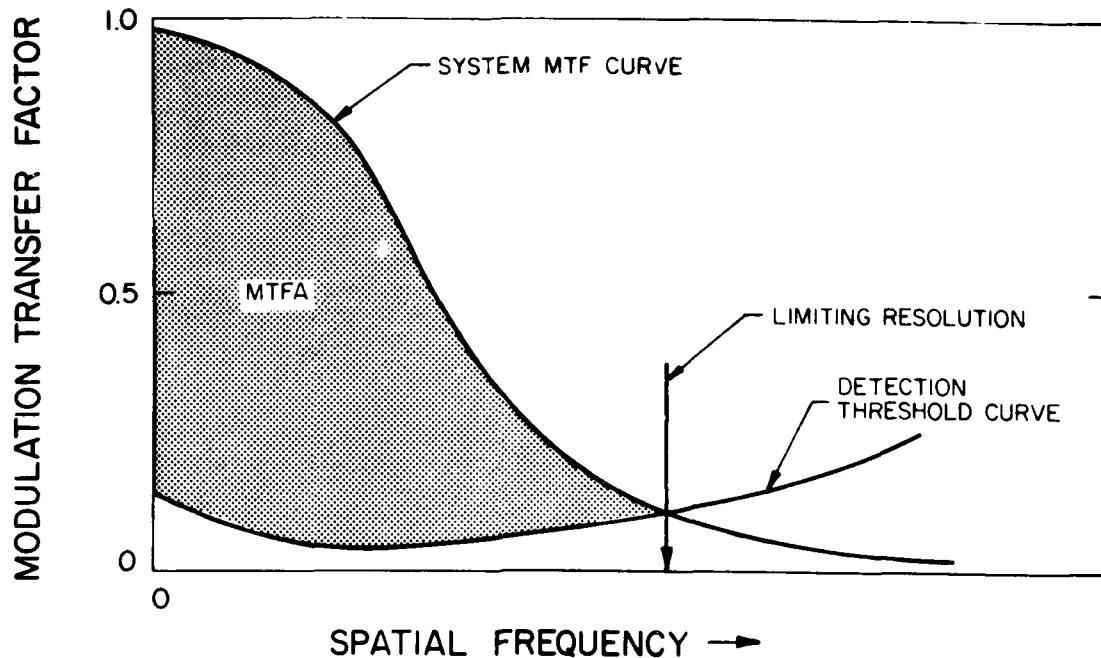


Figure 22. MTF concept.

van Meeteren suggested that the ICS is more sensitive to small changes in the shape of either the MTF of the system or the CTF than would be the MTFA and is therefore more sensitive to small changes in image quality.

Subjective Qualities Factor (SQF). Granger and Cupery (1972) have defined the SQF as being based upon the MTF of the system in conjunction with the contrast sensitivity function. Using the CSF function of Schade (1964, cited by Snyder, 1980), which shows the major sensitivity to lie between 10 and 40 lines per millimeter at the retina, Granger and Cupery defined the SQF for photographic images as the integral of the system MTF between the limits of 10 and 40 cyc/mm when the MTF has been scaled to the retina of the observer by appropriate considerations of the magnification of the system. The SQF is defined mathematically as

$$SQF = k \int_{f=10}^{f=40} |R(f)| d(\log f), \quad (34)$$

where $R(f)$ = the optical transfer function, f = spatial frequency, and k = a normalizing constant.

Perceived Modulation Quotient (PMQ). This metric is an extension of van Meeteren's ICS metric. The difference is that the CTF is divided into the system MTF. The values from this metric are defined on an absolute scale and the ICS metric values are normalized. These absolute values may be important for some applications (Beaton, 1984). It is defined as

$$PMQ = \Delta\omega\Delta\nu \sum \sum \{T_s(\omega, \nu)/T_e(\omega, \nu)\}. \quad (35)$$

Visual Capacity (VC). Cohen and Gorog (1974) took yet another approach to the modification of the MTF concept and built upon Schade's EP metric, extending it to a more modern knowledge of visual perception. In this approach, the visual capacity (VC) metric is defined as,

$$VC = A\Delta\omega\Delta\nu \sum \sum T_s(\omega, \nu)^2 T_e(\omega, \nu)^2, \quad (36)$$

where A denotes the area of the display device. The rationale behind this metric is that the EP is related to the width of edge transitions (sharpness) in the image field and that VC must therefore express the perceptual width of these edge transitions. Normalizing the summed quantity (perceived

edge transitions) by the area of the display is suggested as a means of expressing the maximum number of perceived edge transitions within the image.

Discriminable Difference Diagrams (DDD). Subsequent work by Carlson and Cohen (1978) built upon the earlier RCA activity, and developed a model to predict the just-noticeable differences in contrast discrimination for sine-wave gratings. Using the concept of independent spatial frequency channels in the visual system, these researchers have developed a series of discriminable difference diagrams (DDDs) which correspond to a variety of display conditions. A DDD indicates the increases in modulation necessary to achieve a just-noticeable difference in modulation as a function of frequency. Vertical lines are centered at each spatial frequency channel, and small tick marks indicate the increments at each just-noticeable difference. The number of just-noticeable differences reflects the perceptual extent of image structure in each spatial frequency channel, limited only by the MTF of the display system at that channel. Thus, an image quality metric derived from this approach is the sum of the just-noticeable differences under the MTF, given by

$$JND = \sum_{i=1}^N J(i), \quad (37)$$

in which $J(i)$ indicates the number of just-noticeable differences at channel i , with summation over the N channels.

Just-Noticeable-Difference-Area (JNDA). Task (1979) proposed a metric which he termed the Just-Noticeable-Difference-Area (JNDA) as a possible means of linearizing the contrast axis of the MTF metric. It is defined by transforming the display system MTF curves to JND levels using the DDDs of Carlson and Cohen (1978), then integrating to find the area under the resulting curve. The effect of this transformation is to weight the lower contrast levels more heavily than the higher contrast levels.

Displayed Signal-to-Noise Ratio. Using the analysis of Schade (1953) as a background, Rosell (1971) developed an approach for analyzing television systems which takes into account the temporal and spatial integration capability of the visual system. Rosell's approach is to relate all system parameters to the analytically derived SNR_D . Assuming that the human observer required an SNR_D of approximately 2.8 for a 50% probability of detection, system trade-offs can be made to achieve this or some other level of detection through the relationship between detection probability and SNR_D . Many laboratory studies have been performed to establish this probability of detection as a function of size for geometric figures and single tactical vehicles. Observer confidence levels, task loading, ambient environments, dynamic scenes, target textural characteristics, and other factors have not been considered.

There are many variants of the SNR_D concept, depending on whether one assumes the limitations in the line-scan system to be photon limited, preamplifier limited, display limited, etc. For purposes of discussion, however, an elementary calculational formula is given by Rosell and Willson (1973):

$$SNR_D = [(a/A)t\Delta f_v]^{0.5} SNR_V \quad (38)$$

in which SNR_D = displayed signal-to-noise ratio; a = the area subtended by the target at the display; A = total display area; t = the integration time of the eye, assumed to be constant at 0.2 s, Δf_v = video bandwidth, in MHz; and SNR_V = signal-to-noise ratio in the video, defined as the peak to peak signal divided by RMS noise.

The key to the SNR_D concept lies in the bracketed term in this equation. Essentially, this term provides for both spatial and temporal integration of the signal, and reflects the visual system's spatial and temporal integration capacities. The larger the portion of the display subtended by the target, the greater the signal, with signal strength directly proportional to the square root of the target area, a . In addition, the signal is integrated over the integration time of the visual system, t , which is assumed to be a constant, 0.2 s. More recently, Almagor, Farley, and Snyder (1979) have shown that the integration time is decidedly not constant and varies greatly with adapting luminance, individual observer differences, and the noise level of the display. In fact, these investigators have shown that the visual system typically trades off spatial integration with temporal integration to obtain an optimum visual image.

Hufnagel's Q3. Hufnagel (1965) proposed a system quality metric to account directly for system noise levels in addition to system resolution. As described by Beaton (1984), Hufnagel's metric uses the noise spectral density or Wiener noise power spectrum given by

$$W(\omega, \nu) = \langle \Delta\omega\Delta\nu \sum \sum D(x,y) \exp\{-j2\pi(\omega x + \nu y)\} \rangle^2, \quad (39)$$

where $W(\omega, \nu)$ refers to the two-dimensional Wiener spectrum, the symbol $\langle \rangle$ denotes an ensemble average, and $D(x,y) = L(x,y) - \langle L(x,y) \rangle$ represents the deviations in intensity relative to the mean level. Hufnagel defined the Q3 metric as

$$Q3 = \frac{\Delta\omega\Delta\nu \sum \sum T_s(\omega, \nu)^2 T_e(\omega, \nu)^2}{1 + \{k\Delta\omega\Delta\nu \sum \sum W(\omega, \nu) T_e(\omega, \nu)^2\}} \quad (40)$$

where k is an arbitrary scaling constant. The Q3 resembles a signal-power-to-noise-power ratio.

Signal-to-Noise (SN). Beaton (1984) pointed out that one problem with the Q3 metric is that the scaling constant k must be evaluated, post hoc, to assess the correlation with human performance levels. From previous work showing that noise has a large effect on human quality judgements, it is assumed that k is much greater than one

since the volume under the displayed noise spectrum is much less than unity. Another signal-to-noise (SN) ratio was defined by Beaton, which does not include the experimental constant:

$$SN = \frac{\Delta\omega\Delta\nu \sum \sum T_s(\omega, \nu)^2 T_e(\omega, \nu)^2}{\{\Delta\omega\Delta\nu \sum \sum W_i(\omega, \nu) T_e(\omega, \nu)^2\}^{0.5}} \quad (41)$$

where the denominator represents the root mean square (RMS) deviation of the perceptually weighted noise signal.

Visual Efficiency (VE). Overington (1976, 1982) has developed a sophisticated mathematical model of human visual performance for simple and complex visual environments, basing much of the approach upon basic mechanisms in visual perception. In developing the model, Overington assumes that the illumination gradients between retinal photoreceptors provide important information for target detection, and uses the derivative of the edge (line) spread function (or the Fourier transform) to obtain the following metric, which assumes that the photoreceptor spacing is 25 arcmin:

$$VE = \frac{\sum \sum [T_s(\omega, \nu) T_e(\omega, \nu)] \cos[2\pi(\omega x/N)] \cos[2\pi(\nu y/N)]}{\sum \sum T_e(\omega, \nu) \cos[2\pi(\omega x/N)] \cos[2\pi(\nu y/N)]} \quad (42)$$

in which $x = y = 25$ arcmin.

When $VE > 1$, the perception of image detail is limited by the optics of the eye, whereas when $VE < 1$, edge transitions are limited by the sharpness of the image. Overington (1975) suggests that, in the absence of empirical performance data, the VE metric contains the same fundamental information as the MTF-type metric and therefore should yield similar correlations with performance.

Information Content (IC). The concept of information theory (Shannon and Weaver, 1949) has had a noticeable impact upon developments in image quality metrics, as it has in other technical areas. As applied to images, the amount of information (in bits) in an image is:

$$IC = N \log_2(L), \quad (43)$$

in which IC = image information, in bits; N = number of pixels in an image; and L = number of response levels.

Schindler (1976, 1979) has considered in detail the application of IC to pictorial displays and has derived an equivalent spatial frequency expression for information content, given by

$$IC = \Delta\omega\Delta\nu \sum \sum \log_2 \left[1 + \frac{T_s(\omega, \nu)}{T_d(\omega, \nu)} \right], \quad (44)$$

in which $T_d(\omega, \nu)$ refers to the "just-detectable" response level of the imaging system. The IC metric has units of bits per spatial frequency. Beaton (1984) used the same metric, except that he substituted $T_e(\omega, \nu)$ for $T_d(\omega, \nu)$.

Pixel Error Measures

The MTF measures of image quality discussed above do not define objectively the quality of the content of any particular image. Instead they provide a measure of the extent to which a signal is transmitted through a system regardless of the content of that signal. In contrast to this, pixel error measures of image quality are "image dependent" measures. These measures are based on an error or variance concept in which the extent of the difference in intensity levels, averaged in some fashion across pixels, is taken as a measure of the degradation of an image between the original image and the image whose quality is being measured.

All of the pixel error metrics of interest perform similar calculations on the x, y image arrays, essentially determining the differences between corresponding pixels in the original and the to-be-evaluated image. These differences are then treated mathematically in some fashion to result in a summed or multiplied term which serves as an overall index of quality.

Pixel error metrics of image quality are less supported by empirical vision research than are the MTF-based metrics. While some authors of pixel error metrics claim a "good physical and theoretical basis" to vision (e.g., Granrath, 1981), it can be argued that the correspondence is not well substantiated, at least to the satisfaction of the visual science community.

Normalized Mean Square Error (MSE)

This metric is the basic quantity from which most of the other pixel-error metrics are derived or borrowed. It is defined as

$$\text{MSE} = \frac{\Delta\omega\Delta\nu \sum \sum [M_o(\omega, \nu) - M_m(\omega, \nu)]^2}{\Delta\omega\Delta\nu \sum \sum M_o(\omega, \nu)^2}, \quad (45)$$

in which $M_o(\omega, \nu)$ and $M_m(\omega, \nu)$ refer to the modulation spectra of the original and modified images.

This equation, in its basic form, is simply the sum of the normalized squared deviations between the two images, with the summation unweighted over all pixels. Variations of this general concept have been created by the application of different weighting functions (Pratt, 1978). Four of these weighted approaches follow.

Point Squared Error (PSE)

The PSE normalizes the squared deviations with respect to the maximum value of the original distribution, as given by

$$\text{PSE} = \frac{\Delta\omega\Delta\nu \sum \sum \{M_o(\omega, \nu) - M_m(\omega, \nu)\}^2}{\max\{M_o(\omega, \nu)^2\}}. \quad (46)$$

Perceptual Mean Square (PMSE)

This metric weights the deviations in the MSE by the MTF of the visual system, and is given by

$$\text{PMSE} = \frac{\Delta\omega\Delta\nu \sum \sum \{C_i(\omega, \nu)[M_o(\omega, \nu) - M_m(\omega, \nu)]\}^2}{\Delta\omega\Delta\nu \sum \sum C_i(\omega, \nu)^2 M_o(\omega, \nu)^2}. \quad (47)$$

Image Fidelity (IF)

Linfoot (1960) suggested that the MSE, with appropriate normalization, may be interpreted as a fidelity deficit in the modified image as compared to the original image. He defined the IF metric as unity minus the fidelity deficit, or

$$\text{IF} = 1 - \frac{\Delta\omega\Delta\nu \sum \sum \{M_o(\omega, \nu) - M_m(\omega, \nu)\}^2}{\Delta\omega\Delta\nu \sum \sum M_o(\omega, \nu)^2}. \quad (48)$$

He also suggested two other variants of the MSE which use different normalization values. Their mathematical descriptions follow.

Structural Content (SC)

Structural content is defined as

$$SC = \frac{\Delta\omega\Delta\nu \sum \sum M_m(\omega, \nu)^2}{\Delta\omega\Delta\nu \sum \sum M_o(\omega, \nu)^2} \quad (49)$$

Correlational Quality (CQ)

The CQ metric is defined by the following equation:

$$CQ = \frac{\Delta\omega\Delta\nu \sum \sum M_o(\omega, \nu)M_m(\omega, \nu)}{\Delta\omega\Delta\nu \sum \sum M_o(\omega, \nu)^2} \quad (50)$$

As pointed out by Beaton (1984), there are some interesting relationships among the various metrics. For example, SC may be interpreted as a normalized equivalent of EP. Since EP is related to the width of edge transitions, the SC metric expresses the width of edge transitions in the modified image normalized with respect to the original image. In addition, SC retains the basic form of the Strehl Intensity Ratio if the original image is assumed to be the equivalent of an aberration-free image. The QC metric can be interpreted as the cross correlation of the original image with the modified image, normalized to the original image.

In many respects, these pixel error concepts are simply discrete calculational formulae, for digitized images, of the continuous image concepts advanced under the MTF-based measures. For this reason, it is not surprising that similar correlations have been found for the various measures with observer performance.

In computing the correlations between observer performance and the various image quality metrics, many times scatter plots are made of the data. This is done in order to determine visually the relationship between the two variables. If a linear relationship is not apparent, one or both of the variables of interest may be transformed in some way in order to provide the highest correlation between the two variables. Usually some logarithmic transformation of the data is carried out. This will become apparent in the next section.

Behavioral Validation of Image Quality Metrics

The above section describes several image quality metrics, their associated formulae, and some of the limitations and assumptions of each. This section describes those metrics which have been behaviorally validated, that is, those which have been related experimentally to observer performance. As stated in the introduction to Section 4, such validated measures of image quality can be used to develop models which can predict an observer's performance in terms of information extraction from displays. One of the primary goals of the present research is the development of such models with the specific application to matrix-addressed displays and their associated failure modes. In many cases, the metrics proposed by theorists have never been (or rarely have been) subjected to experimental validation.

The major method used for validation has been to alter the display in some manner, such as changing the system MTF by adding blur or noise to the system, to produce different levels of the image quality metric of interest. These different levels are then correlated with the performance of observers while they view the displays of differing quality. Performance measures which have been used are information extraction, subjective ranking of the quality of image displayed, proportion of correct responses, response time, and search time, among others. Depending on how well these measures correlate, equations can be developed to predict the observer's performance given the value of the image quality metric.

Unfortunately, many times cross-study comparisons of the metrics are virtually impossible. For example, variations in specific design parameters or in the techniques of synthetically manipulating image quality are incompletely controlled, resulting in indeterminate concomitant variation in other potentially relevant factors.

However, it is still instructive to compare the studies which have attempted to validate the image quality metrics. Although an absolute comparison of results cannot be made, much can be learned by such a comparison. For example, the metrics which have received the greatest emphasis in research are revealed. Also, the types of displays studied, the kinds of imagery used, and the performance measures of observer performance used in past research can be learned. By evaluating the past research in such a manner, the methods which have been successful in past studies and the areas in need of research are revealed. This approach helps to determine the important research topics for the present study and lays the foundation for the methods used to obtain the goals of this research.

Monochrome Displays

The studies which have related image quality metrics to observer performance from monochrome displays are listed in Table 22. The metrics of study, the performance measures used, the kinds of displays and imagery used, and the correlations obtained between the metrics and observer performance are summarized in Tables 23 - 26. The reference number in each table corresponds to the appropriate reference in Table 22.

Table 22

References Which Relate Image Quality Measures to Observer Performance

-
1. Beaton, 1984
 2. Blumenthal and Campana, 1981
 3. Borough, Fallis, Warnock, and Britt, 1967
 4. Granger, 1974
 5. Granger and Cupery, 1972
 6. Gutmann, Snyder, Farley, and Evans, 1979
 7. Hall, 1981
 8. Klingberg, Elworth, and Filleau, 1970
 9. Snyder, 1974
 10. Snyder, 1976
 11. Snyder, 1967
 12. Snyder and Maddox, 1978
 13. Snyder, Keesee, Beamon, and Aschenbach, 1974
 14. Snyder, Beamon, Gutmann, and Dunsker, 1980
 15. Task, 1979
-

Table 23 shows the correlations between the MTF-based image quality metrics and observer performance for film displays using literal imagery. From this table, it is obvious that not all the MTF-based metrics have been studied in this context, and that those which have been studied provide very good correlations with different performance measures. Studies emphasizing the subjective ranking of imagery and the MTF metric have been emphasized the most. However, other measures of performance have correlated well with the various MTF-based metrics (reference 15).

The film displays represent continuous tone imagery. Although the matrix-addressable displays in the current program are spatially discrete displays, there are certain arrangements of pixel size and viewing distance which do not allow the observer to discern the edges between the pixel elements. Under these conditions, for all practical purposes the discrete display becomes a continuous one to the observer. In this instance the metrics which have shown promise in Table 22 may be applied.

Table 24 shows the correlations between the MTF-based metrics and observer performance for CRT displays using literal imagery. From these tables, it is apparent that more of the MTF-based image quality metrics have been studied in this context than with film displays and that the MTF metric has received the most attention. Also, the number of performance measures used has increased.

Table 23

Correlations Between MTF-Based Image Quality Metrics and Observer Performance for Film Displays with Literal Imagery

Reference/Performance Measure	Image Quality Metric							
	SNR _D	MTFA	SQF	JNDA	GSFP	ICS	NE	
2. Subjective ranking	.987							
3. Subjective ranking		.92						
4. Subjective ranking			.988					
5. Subjective ranking			.99					
8. Subjective ranking		.97						
8. Information extraction (number of errors)		-.93						
15. Angle subtended at recognition		-.95 ^a	-.979	-.984 ^a	-.858 ^a	-.978	-.729	

^a Log transformation

Table 24

Correlations Between MTF-Based Image Quality Metrics and Observer Performance for CRT Displays with Literal Imagery

Reference/Performance Measure	Image Quality Metric									
	MTFA	GSFP	ICS	PMQ	VC	Q3	SN	IC	EP	
1. Information extraction ^b	.79	.80	.95	.95	.87	.87	.95	.86	.78	
1. Subjective ranking ^b	.71	.73	.95	.95	.90	.90	.95	.84	.69	
6. Proportion of correct Rs	.593									
6. Slant range at recognition	.73									
9. Proportion of correct Rs	.87 ^a									
9. Response time	-.92 ^a									
10. Proportion of correct Rs	.211									
10. Slant range at recognition	.699									
11. Proportion of correct Rs	.97									
11. Slant range at recognition	.76									
13. Proportion of correct Rs	.965									

^a Log transformation

^b Digitally addressed CRT

Table 24 (continued)

Correlations Between MTF-Based Image Quality Metrics and Observer Performance for CRT Displays with Literal Imagery

Reference/Performance Measure	Image Quality Metric							
	MTFA	SNR _D	SQF	ICS	JNDA	EP	GSFP	
13. Proportion of incorrect Rs	-.973	-.817						
13. Slant range at recognition	.765	.514						
13. Proportion of correct Rs		.968						
14. Proportion of correct Rs	.473							
15. Angle subtended at recognition	-.878 ^a		-.781	-.818	-.906 ^a	-.726	-.847 ^a	
15. Slant range at detection	.866 ^a		.803	.837	.902 ^a	.761	.869	

^a Log transformation

AD-A180 272

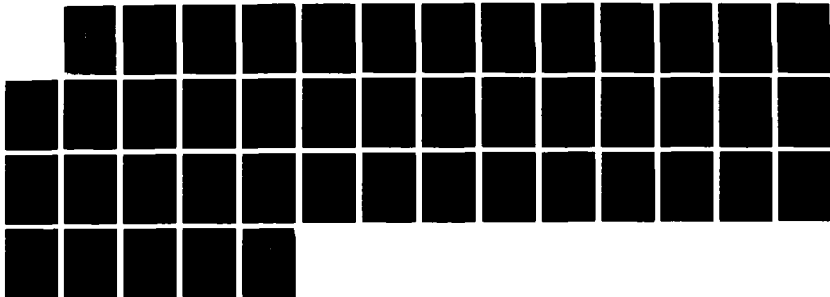
A LITERATURE REVIEW AND EXPERIMENTAL PLAN FOR RESEARCH
ON THE DISPLAY OF.. (U) VIRGINIA POLYTECHNIC INST AND
STATE UNIV BLACKSBURG HUMAN FAC. . . J J DECKER ET AL.
FEB 87 MEL-TN-4-87 DAAA15-85-R-0066

2/2

UNCLASSIFIED

F/G 9/5

NL



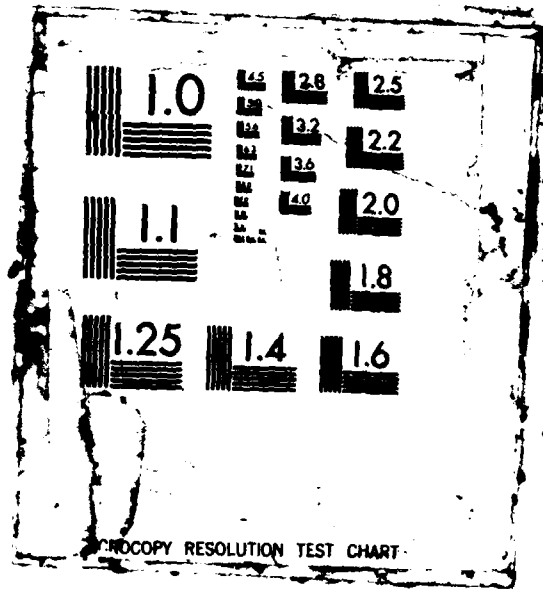


Table 24 (continued)
 Correlations Between MTF-Based Image Quality Metrics and
 Observer Performance for CRT Displays with Literal Imagery

Reference/Performance Measure	Image Quality Metric							
	IR	SSF	EW	PEP	PIR	PSSF	PEW	
1. Information extraction ^b	.78	.79	.78	.87	.95	.91	.95	
1. Subjective ranking ^b	.71	.71	.71	.89	.95	.89	.95	

^b Digitally addressed CRT

The performance measures most important to the present study are information extraction, proportion of correct responses, and perhaps, for photointerpretation tasks, slant range at recognition. These measures correlate from good (references 1, 9, 11, 13) to fair (references 6, 10, 14) with the different image quality metrics.

Of particular importance to the present study are the results from reference 1 (Beaton, 1984). This study used a photointerpretation task with digitally addressed imagery displayed on a CRT. The images were degraded by varying levels of noise and blur which produced varying values of each of the image quality metrics. Some of the correlations between the performance measures of information extraction and image quality metrics, and between subjective ranking of images and image quality metrics are very good. Among the metrics showing the greatest degree of relatedness to performance are ICS, PMQ, SN, PIR, and PEW.

Beaton's study is important to the present study in at least two ways. One, it shows that the MTF-based metrics can be successfully applied to displays producing digitally-addressed imagery. Next, it reveals that the much researched MTF metric does not always produce the best correlation with observer performance. For the present study this means that there are valid reasons to apply the MTF-based metrics to evaluate the image quality of digital displays and that other metrics besides just the MTF metric should be studied.

Table 25 shows the correlations between the pixel-error image quality metrics and observer performance for CRT displays presenting digitally addressed literal imagery. It is evident that not much behavioral research has been performed to validate these metrics. This is somewhat disappointing because the pixel-error metrics represent metrics which may be able to account for the typical failure modes of matrix-addressed displays, largely because the pixel-error metrics are "image dependent" metrics in that they are used to compare an original image with a degraded image. The degraded image in this case would be one in which some cells or lines have failed. Because of the importance of understanding the effect of the failure modes on display quality, emphasis should be placed on extending the behavioral validation of the pixel-error metrics.

Table 25

Correlations Between Pixel Error Image Quality Metrics and Observer Performance for CRT Displays with Literal Imagery

Reference/Performance Measure	Image Quality Metric				
	.hr std left right				
	NMSE	PMSE	IF	SC	CQ
1. Subjective ranking ^a	.60	.15	.60	.19	.11
7. Subjective ranking ^a	.85	.92			

^a Digitally addressed CRT

Table 26 shows the correlations between the MTF metric and observer performance for CRT displays with alphanumerics. The reference of most importance is number 12 (Snyder & Maddox, 1978). In this study, Snyder and Maddox used dot matrix characters and showed that $r = .82$ between the proportion of correct responses and the MTF metric. Of more importance is their development of an empirical image quality model dealing with the prediction of observer performance from displays presenting dot-matrix alphanumerics.

All of the above metrics of image quality have one thing in common. They are based upon some theoretical approach to the notion of image quality and the quantification of the visual system, and lead directly to a model of image quality based upon that theoretical approach. A totally different approach was taken by Snyder and Maddox (1978). They offered no pet theory or concept and determined empirically which concepts predict observer performance, letting the resulting pool of predictors define quantitatively what is meant by "image quality."

Table 26

**Correlations Between the MTFA Image Quality Metric
and Observer Performance for CRT Displays with Alphanumerics**

Reference/Performance Measure	MTFA Image Quality Metric
6. Proportion of correct Rs	.349
6. Search time	-.703
12. Proportion of correct Rs ^b	.82 ^a
14. Search time	-.88

^a Log transformation

^bDigitally addressed CRT

Using three different tasks, they performed experiments which varied the structure of the display in terms of pixel size, shape, contrast, spacing, and the like. They measured observer performance on two different search tasks and a reading task and correlated these performance measures with a variety of physical measurements of geometric and photometric characteristics of the image. Table 27 lists the predictor variables which were tested in a stepwise, linear multiple regression approach. In this statistical approach, all known variables are permitted to enter into a linear prediction equation, and the computed result is a "model" that defines the best predictive combination of any or all of the variables. The resulting R^2 value gives the percent of the variation among the various display conditions that can be predicted by the model.

Table 28 indicates the resulting prediction equations from the Snyder and Maddox experiments for two of the tasks, the reading task and the structured visual search task. It can be seen that this empirical model, which has subsequently been cross-validated, predicts 50% of the variance for the search task and 52% of the variance in the reading task. Of perhaps more interest are the combinations of variables which entered into the prediction equations. These predictor variables are almost entirely modulation and MTFA type measures, and generally support the results which have been previously obtained for these types of image quality measures.

As noted by Snyder and Maddox (1978), the equations in Table 28 represent the best empirically derived measures of image quality for digitally addressed displays, for the purpose of design specification. They do not deal directly with the recommended dynamic range of a given image or any other image-specific parameters as do some of the measures described above. Thus, these equations are useful by the designer to optimize displays particularly for the presentation of alphanumeric information. This research was done with digitally addressed CRT displays.

Given the success of the model developed by Snyder and Maddox (1978) for alphanumerics, such an approach may be the best way to predict performance from matrix addressed displays given their failure modes. Their method should be extended to the study of literal images and graphics as well as alphanumerics.

Multichromatic Displays

In spite of the plethora of research and modeling devoted to monochromatic displays, there appears to be no accepted metric of image quality devoted to multichromatic displays. Some metrics have been proposed which have been derived from prior "monochrome" metrics, such as the least-squared deviations from "true" color (Pratt, 1978). Other studies have derived three optical transfer functions corresponding to the three tristimulus values (Bescos and Santamaria, 1977). Such an approach could possibly lead to the development of an MTF-based color image quality metric.

Some of the above mentioned metrics have been applied or suggested as being useful as color image quality metrics. For example, Granger (1974) found that when the SQF metric was modified to include the spectral luminosity response of the visual system, it was able to accurately predict the image quality rank of color pictures. Also, Overington (1976) discussed the fact that the VE metric may be compatible with modern models of color vision.

Table 27

Pool of Predictor Variables, from Snyder and Maddox (1978)

Vertical	Horizontal	Description
VFREQ	HFREQ	Fundamental spatial frequency (cyc/deg)
VFLOG	HFLOG	Base 10 log of fundamental spatial frequency
VSQR	HSQR	Square of (fundamental spatial frequency minus 14.0)
VMOD	HMOD	Modulation of fundamental spatial frequency
VDIV	HDIV	Fundamental spatial frequency divided by modulation
VLOG	HLOG	Base 10 log of VDIV and HDIV
VMTFA	HMTFA	Pseudo-modulation transfer function area
VMLOG	HMLOG	Base 10 log of VMTFA and HMTFA
MCROS	HCROS	Spatial frequency at which modulation curve crosses the threshold curve
VRANG	HRANG	Crossover frequency minus fundamental frequency

There have been studies of the influence of individual chromatic display parameters on both subjective quality estimates and upon some simple observer performance, as noted in Section 3. But apparently no effort has been made to develop an all-inclusive model of all the chrominance and luminance variables in a complex display. Although there have been efforts to develop a valid metric of color contrast (e.g., Post, Costanza, and Lippert, 1982), there does not seem to exist a valid metric of color image quality. Such a metric needs to be developed, not only for discrete pixel displays but also for continuous image displays.

Areas in Need of Research

Several areas in need of research can be determined by the review of the image quality metrics. Those which apply to the present program are summarized as follows.

1. It is obvious that none of the image quality metrics reviewed in this report have been applied in an attempt to account for the effects of the failure modes of flat-panel display devices on displayed image quality. Therefore, there is no way of predicting the effects of such failures on observer performance given objective measures of image quality. Due to the absence of such studies, the importance of the present program is further emphasized.

One of the primary purposes of the present research is to develop or adapt an image quality metric (or metrics) which takes into account the failure modes of displays when defining the objective image quality. The MTF metrics defined above provide a measure of the extent to which a signal is transmitted through a system regardless of the content of that signal. Because of this, such measures do not represent what happens to the displayed modulation (or image quality) when a cell or line fails. Pixel error measures of image quality, on the other hand, can be used to evaluate

Table 28

Extended Predictive Equations, from Snyder and Maddox (1978)

Task	Metric and Related Information
Tinker SOR	$\text{Adjusted Reading Time (s)} = 5.74 + 0.3111(\text{HFREQ})$ $+ 2.379(\text{HMOD}) + 4.365(\text{HLOG})$ $- 14.973(\text{HFLOG}) + 1.112(\text{VMLOG})$ <p>Correlation Coefficient $R = 0.72$</p> <p>$R^2 = 0.525$</p> <p>Asymptotic $R^2 = 0.637$</p>
Menu Search	$\text{Search Time (s)} = 7.27 + 0.027(\text{HDIV}) + 2.159(\text{HLOG})$ $+ 5.916(\text{VFLOG}) - 0.339(\text{VMTFA})$ $- 0.054(\text{VRANG}) + 5.487(\text{VMLOG})$ <p>Correlation Coefficient $R = 0.71$</p> <p>$R^2 = 0.500$</p> <p>Asymptotic $R^2 = 0.575$</p>

the displayed image when there are cell/line failures. As described previously, such metrics can be used to relate the intensity distributions of an original image to a degraded version of the image such that the differences in the statistical intensity distributions are a measure of the degradation of the reduced image quality. However, there is a problem associated with the pixel error measures which limits their generalizability: they are "image dependent." That is, they can only be used to compare the image quality variations of a given image and do not provide a measure of quality regardless of the image content. Since many different kinds of images can and need to be displayed, this is a serious limitation.

The solution may lie in developing an empirical image quality model following the method of Snyder and Maddox (1978). Among the predictor variables included in such a model could be some representation of the percentage of cells or lines which have failed.

It is apparent then from the review of image quality metrics that the development of a metric or model which takes into consideration the failure modes of matrix addressable displays is an area in need of research.

2. From reviewing the conditions under which the various image quality metrics have been studied, there is a complete lack of the use of cartographic/symbolic display content. Much emphasis has been placed on the use of literal imagery, especially in the context of photointerpretation. Some studies have also dealt with alphanumeric. The lack of data concerning cartographic and symbolic information shows both areas are in need of research.

3. As already mentioned in a previous section, there is no image quality metric available for the objective evaluation of multichromatic displays. With the recent development and foreseeable implementation in the field of full-color solid-state matrix-addressed displays, the problem becomes more severe in determining the level of display failure that will affect the utility of the display under operational conditions. It is therefore important to develop some sort of multicolor-display image quality metric which could be used for display evaluation, user performance prediction, and device quality assurance as well as refining the monochrome image quality metrics for the evaluation of monochrome display content.

4. Very few of the experiments which have attempted to behaviorally validate the image quality metrics have used digitally addressed displays. In addition, only Snyder and Maddox (1978) have used actual flat-panel display devices to assess image quality. In that study, such devices were used to provide a validation of the empirically derived model, the model itself having been developed using a CRT to simulate the parameters of the actual flat-panel displays. Because of the lack of

use of actual flat-panel displays, it is important that any metric or model developed in the present program should be validated on the actual flat-panel displays of interest. This approach will provide the additional data needed to properly refine the metrics or models so as to achieve maximum prediction of performance.

SECTION 5

EXPERIMENTAL RESEARCH PLAN

The following research plan is divided into four major sections: (1) task descriptions and dependent measures; (2) monochromatic display research; (3) multichromatic display research; and (4) quality metrics. Each of the monochromatic and multichromatic sections is divided further into studies dealing with alphanumeric and cartographic/symbolics. The major difference between these two subdivisions is that the cartographic/symbolic studies will have a map background on which both alphanumeric and symbolic information will be overlaid. By studying the effects of cartographic/symbolics in this manner, comparisons across studies can be made (i.e., comparisons between selected alphanumeric experiments and cartographic/symbolic experiments) to determine the effects of the complex background.

It was decided that studies dealing with literal imagery would not be included in this experimental plan. The reason for this is due to the limiting aspects of current and near-term flat-panel display technologies. Currently, such devices have an inadequate gray scale for the presentation of literal images. Also, although multichromatic flat-panel displays are under development (e.g., thin-film EL), they are limited in their ability to display quality full color literal imagery.

Task Descriptions and Dependent Variables

Two tasks will be used to investigate the effects that various flat-panel display parameters have on performance. Random search tasks and information extraction tasks have been used successfully in past research concerning the legibility and readability of dot-matrix displays (Abramson and Snyder, 1984; Albert, 1975; Snyder & Maddox, 1978). For both tasks, dependent measures of response speed and accuracy will be taken.

The random search task allows for comparison of several independent variables across the alphanumeric and cartographic/symbolic studies. With the random search task, single nonoverlapping characters are positioned randomly on the display. The target character is displayed top center on the display. The subject simply has to find the target character among the randomly displayed characters and give its location. By using the same task for both the alphanumeric and cartographic/symbolic studies the effects of the complex background can be determined.

By design, the information extraction tasks must be different for the alphanumeric and cartographic/symbolic experiments. For the alphanumeric experiments, a modification of the Tinker Speed of Reading Test will be used. This test has been shown to be an accurate and reliable measure of operator reading performance with electronic displays (Burnette, 1976; Snyder & Maddox, 1978). For the cartographic/symbolic experiments the information extraction task will involve searching the display and interpreting the displayed symbols and cartographics. The precise form of this task will be developed subsequently.

Monochromatic Display Research

Research investigating the effects of display variables on user performance with monochromatic displays is necessary for two reasons. In many tasks color coding of information is not required; therefore, many future displays will be monochromatic. Also, many flat-panel displays will not have full color capabilities for many years. Most of the existing research has been on monochromatic displays; however, a great deal of additional data is still needed. These data can be used to develop and recommend metrics that predict information extraction performance.

Several variables have been selected for investigation based on:

- previous research which illustrated that the display parameter(s) are critical to user performance (e.g., contrast),
- lack of previous research investigating several of the variables or their interactions with other variables,
- their potential for user performance prediction,
- the belief that the variables may have differential effects for

- different types of tasks, and
- the likelihood of important interactions with other variables of interest.

The variables selected include:

- matrix/character size
- contrast (character or background luminance)
- character font
- case
- polarity
- symbol height/width
- symbol orientation
- display failure (type, mode, and percent failed)
- element size, shape, and spacing
- uniformity
- background clutter

As the research progresses new variables may also be added based on the experimental findings. Alphanumeric and cartographic/symbolic experiments have been partially defined using the variables listed above. The purpose, tasks, and variables of each experiment are briefly outlined below.

Monochromatic Alphanumeric Research

Experiment 1: Optimum Character Study

Purpose: To determine an optimal character set under nondegraded conditions. The effects of degradation on this optimal set (and perhaps other less optimal sets) will be assessed in future experiments.

Task(s): Information extraction and random search

Variables:

- matrix/character size (7 x 9, 9 x 11, 11 x 15)
- contrast (3:1, 6:1, 10:1)
- font (3 types)
- case (upper and mixed)
- polarity (positive, negative)

Matrix/character size, font, and case are character definition variables which have been found to interact with one another in previous research. Contrast, a critical variable to legibility/readability, also has been found to interact with these variables. Case and polarity have not been well researched in the past, but are considered important for defining a character set. Special symbols will not be used in this study; however, they will be investigated in Experiment 2.

Experiment 2: Character Modification

Purpose: This study is an extension of Experiment 1 to define further an optimal character set. This study will include special symbols as well as alphanumerics.

Task(s): Random search

Variables:

- matrix/character size (3, selected from Experiment 1 results)
- symbol height/width (3)
- symbol orientation (6 angles, to be determined by a pilot study)

Contrast, font (for alphanumerics), and polarity will be held constant based on the results of Experiment 1. The literature review pointed out the need for research investigating optimal character heights and widths. Matrix/character size will most likely interact with this variable. Symbol orientation is an important variable because symbols are often rotated when used on cartographic/symbolic displays, and currently very few data are available regarding this variable. The information extraction task will not be used because it is unlikely that reading tasks will have rotated characters.

Experiment 3: Failure Modes I

Purpose: To provide quantitative data on the effects of display failures on user performance. These data will then be used to aid in developing a quality metric that can be used for predicting user performance, display evaluation, and device quality assurance.

Task(s): Information extraction and random search

Variables:

- failure type (vertical or horizontal line failures, cell failures)
- failure mode (failed on or off)
- percent failure (4 percent levels)

- polarity (2)
- matrix/character size (3)

Experiment 4: Failure Modes II

Purpose: This is an extension of Experiment 3; however, a different subset of variables will be under investigation along with display failures. This study will use only a subset of the failure levels expected to be used in Experiment 3.

Task(s): Random search

Variables:

- failure type (vertical and horizontal line failures, cell failures)
- failure mode (failed on or off)
- percent failure (3)
- font (3)
- symbol orientation (3)

Matrix size and polarity will be held constant based on results from Experiment 3.

Experiment 5: Element Characteristics

Purpose: To determine whether different element sizes, shapes, and spacings will be differentially affected by display failures. These data will aid in recommending element configurations which are least sensitive to the display degradation caused by display failures.

Task(s): Information extraction and random search

Variables:

- failure type (vertical and horizontal line failures, cell failures)
- failure mode (failed on or off)
- percent failure (3 levels)
- element size, shape, spacing (27 combinations)

Experiment 6: Uniformity

Purpose: To determine the effects of nonuniformity on user performance.

Task(s): Information extraction

Variables:

- large area uniformity (3 levels, to be determined)
- small area uniformity (3 levels, to be determined)
- edge discontinuities (3 levels, to be determined)

As discussed in the literature review, no data currently exist which illustrate the effects of nonuniformity on user performance. A reading task will be used for this experiment because it is believed that small or large changes in luminance across the display will be apparent in reading or text displays. It is also possible that, like flicker, nonuniformity will not directly affect performance, but it may cause discomfort or user fatigue.

Monochromatic Cartographic/Symbolic Research

Experiment 7: Optimum Character (Symbolic)

Purpose: To determine whether the optimal character sets obtained in the alphanumeric studies can be transferred to tasks with complex backgrounds.

Task(s): Information extraction and random search

Variables:

- matrix/character size (2)
- symbol orientation (4)
- contrast (2)
- background clutter (2)

Only a subset of the levels used in the alphanumeric studies will be used in this study. Specific levels to be investigated will be selected after the completion of Experiments 1 through 4.

Experiment 8: Failure Mode Study (Symbolic)

Purpose: To determine the effects of display failures on user performance with cartographic/symbolic displays.

Task(s): Information extraction and random search

Variables:

- failure type (vertical and horizontal line, cell)
- failure mode (on, off)
- percent failure (4)
- matrix/character size (2)
- symbol orientation (4)

It is likely that display failures will affect cartographic/symbolic displays differently than alphanumeric because of the lack of redundancy in cartographic/symbolic displays, and because more display content interference may exist.

Experiment 9: Element Characteristics (Symbolic)

Purpose: To evaluate the effects of display failures on cartographic/symbolic information displays constructed with different element sizes, shapes, and spacings.

Task(s): Information extraction and random search

Variables:

- element size, shape, and spacing (27 combinations)
- failure type (vertical and horizontal line, cell)
- failure mode (on, off)
- percent failure (3)

Multichromatic Display Research

The purpose of this portion of the research is to investigate the effects of multichromatic display variables on user performance. Many of the display variables used in the monochromatic display experiments will also be used in these experiments to determine whether introducing color differentially affects user performance with those variables. The studies defined in this section are somewhat parallel to the monochromatic display experiments so that the results can be compared among studies. Color contrast metrics will be developed based on these empirical data.

Of primary interest is the effect of display failures on multichromatic information displays. Matrix displays may fail (either off or on) in one, two, or three of the primary colors (assuming a 3-primary display, such as a CRT) causing shifts in hue as well saturation, depending upon the failure type, mode, and original color of the information displayed.

The variables for these studies were selected from the same factors previously listed in the Monochromatic Display Research Section. The variables include:

- background chrominance
- luminance (modulation)
- target chrominance
- matrix/character size
- symbol height/width
- font
- case
- display failures (type, mode, percent failed)
- number of primary colors failed (one, two, or three)
- background clutter
- uniformity

Alphanumeric and cartographic/symbolic studies will be conducted. They are briefly described below.

Multichromatic Alphanumeric Research

Experiment 10: Replication of Color Contrast Experiment

Purpose: This study will partially replicate the studies conducted by Lippert (1984, 1985). Lippert used a task which required subjects to read colored dot matrix numeral strings (targets) against colored backgrounds. This study will replicate several of the conditions; however, different tasks will be used to determine whether Lippert's results will generalize to different tasks.

Task(s): Information extraction and random search

Variables:

- luminance contrast and polarity (7 levels)
- background chrominance (8)
- target chrominance (3)

Experiment 11: Multichromatic Optimum Character Study

Purpose: To evaluate the effects of color on character definition variables to determine whether the optimal character sets defined in the monochromatic studies will transfer to color displays.

Task(s): Information extraction and random search

Variables:

- subset of chrominance/luminance background combinations (3)
- target colors (3)
- matrix/character size (3)

- font (3)
- case (2)

Experiment 12: Multichromatic Failure Mode Study

Purpose: To investigate the effects of display failures on multichromatic alphanumeric displays.

Task(s): Information extraction and random search

Variables:

- subset of chrominance/luminance backgrounds (3)
- failure type (vertical or horizontal line failures, cell failures)
- failure mode (on or off)
- percent failure (3)
- number of primary colors failed (3)
- target color (3)

Experiment 13: Multichromatic Uniformity

Purpose: To evaluate the effects of chrominance and luminance changes across the display on user performance. Each of the three variables listed below will be considered as a separate one variable experiment with five levels per variable with levels to be determined.

Task(s): Information extraction

Variables:

- *Exp. 13a*, small area nonuniformity (5 levels)
- *Exp. 13b*, large area nonuniformity (5 levels)
- *Exp. 13c*, edge discontinuity (5 levels)

Multichromatic Cartographic/Symbolic Research

Experiment 14: Multichromatic Optimum Character Symbolic

Purpose: To determine whether the optimal character sets are differentially affected by complex backgrounds.

Task(s): Information extraction and random search

Variables:

- target color (3)
- matrix/character size (3)
- symbol orientation (6)
- background clutter (2)

Summary

Table 29 is a matrix which summarizes each of the variables included in all of the 14 Experiments previously outlined.

Quality Metrics Analysis

As described in Section 4, a variety of quality metrics exist from previous research on both continuous image and dot-matrix displays. One of the major purposes of the current research is to develop suitable metrics for describing the efficacy and adequacy of matrix addressed displays given their display parameters and failure modes.

Throughout the proposed experiments, radiometric and photometric measures will be taken to carefully define the display characteristics being presented to the experimental observers. At the end of the experimental portion of the research, the quality metrics analysis will combine the results of the previous experimental tasks into an evaluation of suitable metrics for describing matrix addressed displays. Analytical (e.g., correlational) studies will be made of the adequacy of various quality metrics to predict the obtained observer performance results. The prediction accuracy of these models will be reported, as will any new quality metric models which account better for the experimental data.

Table 29

Summary Matrix of Variables In Each Experiment

Variables	Experiment Number													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Matrix/ Character Size	X	X	X				X	X			X			
Contrast (Luminance)	X						X			X				
Font	X			X							X			
Case	X										X			
Polarity	X		X							X				
Symbol height/ Width		X												
Symbol Orientation		X		X			X	X						X
Display Failures (type, mode, % failed)			X	X	X			X	X			X		
Element size, shape, spacing					X					X				
Uniformity						X							X	
Background Clutter							X							X
Background Chrominance										X	X	X		
Target Chrominance										X	X	X		X
Number of Primary Colors Failed												X		

REFERENCES

- Abramson, S. R., & Snyder, H. L. (1984). *Operator performance on flat-panel displays with line and cell failures*(Tech. Report HFL-83-3). Blacksburg: Virginia Polytechnic Institute and State University, Human Factors Laboratory.
- Albert, D. E. (1975). *Prediction of intelligibility of contextual dot matrix characters*. Unpublished Master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Aldersey-Williams, H. (1985, December). LCDs: a bright new face. *High Technology*, 40-42.
- Almagor, M., Farley, W. W., & Snyder, H. L. (1979). *Spatiotemporal integration in the visual system* (Tech. Report AMRL-TR-78-126). Wright-Patterson Air Force Base, OH: Air Force Aerospace Medical research Laboratory.
- Avery, T. E., & Berlin, G. L. (1985). *Interpretation of Aerial Photographs*. Minneapolis, MINN: Burgess, Publishing Co.
- Banks, W. W., Gertman, D. I., & Petersen, R. J. (1982). *Human Engineering Design Considerations for Cathode Ray Tube-Generated Displays*(Tech. Report NUREG/CR-2496). Washington D. C.: U. S. Nuclear Regulatory Commission.
- Bauer, D., & Cavonius, C. R. (1980). Improving the legibility of visual display units through contrast reversal. In E. Grandjean and E. Vigliani (Eds.), *Ergonomic aspects of visual display terminals* (pp. 137-142). London: Taylor and Francis.
- Beamon, W. S. & Snyder, H. L. (1975). *An experimental evaluation of the spot wobble method of suppressing raster structure visibility*(Tech. Report AMRL-TR-75-63). Wright-Patterson Air Force Base, OH: Air Force Aerospace Medical Research Laboratory.
- Beaton, R. J. (1984). *A human-performance based evaluation of quality metrics for hard-copy and soft-copy digital imagery systems*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Bescos, J. & Santamaria, J. (1977). Formation of color images. Optical transfer functions for the tristimulus values. In R. Shaw (Ed.), *SPSE Conference Proceedings* (pp. 38-47). Washington, D. C.: Society of Photographic Scientists.
- Blumenthal, A. H. & Campana, S. B. (1981). An improved electro-optical image quality summary measure. In *Proceedings of the Society of Photo-Optical Instrumentation Engineers, Image Quality, 310*, (pp. 43-52). Bellingham, WA: SPIE.
- Borough, H. C., Fallis, R. F., Warnock, R. H., & Britt, J. H. (1967). *Quantitative determination of image quality* (Tech. Report D2-114058-1). Boeing Company.
- Bracewell, R. (1965). *The Fourier transform and its applications*. New York: McGraw Hill.

- Brown, J. L. (1965). Flicker and intermittent stimulation. In C. H. Graham (Ed.), *Vision and visual perception* (pp. 251- 320). New York: Wiley.
- Buckler, A. T. (1977). *A review of the literature on the legibility of alphanumeric characters on electronic displays*. (TM 16-77). Aberdeen, MD: U.S. Army Human Engineering Laboratory.
- Burnette, J. T. (1976). *Optimal element size-shape-spacing combinations for a 5X7 dot matrix visual display under high and low ambient illumination*. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg VA.
- Campbell, F. W. & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, 197, 551-566.
- Carlson, C. R. & Cohen, R. W. (1978). *Visibility of displayed information: Image descriptors for displays* (Tech. Report ONR-CR213-120-4F). Arlington, VA: Office of Naval Research.
- Carter, R. C. (1982). A design tool for color displays. In *Proceedings of the Human Factors Society 26th Annual Meeting* (pp.589). Santa Monica, CA: Human Factors Society.
- Chao, B. P. (1983). *Human performance evaluations of selected image enhancement/restoration techniques* Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Charman, W. N. & Olin, A. (1965). Tutorial: Image quality criteria for aerial camera systems. *Photographic Science and Engineering*, 9, 385- 397.
- Cohen, R. W. & Gorog, I. (1974). Visual capacity: An image quality descriptor for display evaluation. In *Proceedings of the Society for Information Display*, 15,(pp. 53-62). New York: Palisades Institute
- Cornog, D. Y., & Rose, F. C. (1967). *Legibility of alphanumeric characters and other symbols: II. A reference handbook*. Washington: U. S. Government Printing Office, National Bureau of Standards Miscellaneous Publication 262-2.
- Craford, M. G. (1985). Light-emitting diode displays. In L. E. Tannas (Ed.), *Flat-panel displays and CRTs* (pp. 289- 331). New York: Van Nostrand Reinhold Co.
- Dainty, J. C. & Shaw, R. (1974). *Image science*. New York: Academic Press.
- De Corte, W. (1985, April). High contrast sets of colours for colour CRTs under various conditons of illumination. *Displays*, 95-100.
- Duncan, J., & Konz, S. (1974). Effect of ambient illumination on legibility of displays of liquid-crystals and light-emitting diodes. In *Proceedings of the Human Factors Society 18th Annual Meeting*(pp.102-108). Santa Monica, CA: Human Factors Society.

- Ellis, B., Burrell, G. J., Wharf, J. H., & Hawkins, T. D. F. (1975). The format and colour of small matrix displays for use in high ambient illumination. In *Proceedings of the Society for Information Display*, 16, (pp. 250-258). New York: Palisades Institute.
- Farrell, R. J., & Booth, J. M. (1975). *Design handbook for imagery interpretation equipment* (Tech. Report AD/A-025 453). Boeing Aerospace Company, Seattle, WA: U. S. Department of Commerce National Technical Information Service.
- Gaskill, J. D. (1978). *Linear systems, Fourier transforms, and optics*. New York: Wiley.
- Giddings, B. J. (1972). Alphanumerics for raster displays. *Ergonomics*, 15, 65-72.
- Goodman, L. A. (1974). The relative merits of LEDs and LCDs. In *Proceeding of the Society of Information Display*, 16, (pp.8-15). New York: Palisades Institute.
- Gould, J. D. (1968). Visual factors in the design of computer-controlled CRT displays. *Human Factors*, 10, 359-376.
- Graff, G. (1985, December). Electroluminescence: better picture, higher cost. *High Technology*, 42-45.
- Granger, E. M. (1974, May). Subjective assessment and specification of color image quality. In *Proceedings of the Society of Photo-Optical Instrumentation Engineers*, 46, (pp. 86-91). Bellingham, WA: SPIE.
- Granger, E. M. & Cupery, K. N. (1972). An optical merit function (SQF), which correlates with subjective image judgements. *Photographic Science and Engineering*, 16, 221- 230.
- Granrath, D. J. (1981). The role of human visual models in image processing. In *Proceedings of the IEEE*, 69, (pp. 552-561).
- Gutmann, J. C., Snyder, H. L., Farley, W. W., & Evans, J. E. (1979). *An experimental determination of the effect of image quality on eye movements and search for static and dynamic targets* (Tech. Report AMRL-TR-79-51). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.
- Hall, C. F. (1981, August). Subjective evaluation of a perceptual quality metric. In *Proceedings of the Society for Photographic Instrumentation Engineers, Image Quality*, 310 (pp. 200-204). Bellingham, WA: SPIE.
- Hemingway, J. C., & Erickson, R. A. (1969). Relative effects of raster scan lines and image subtense on symbol legibility on television. *Human Factors*, 11, 331-338.
- Howard, W. E. (1981). Electroluminescent display technologies and their characteristics. In *Proceedings of the Society for Information Display*, 22, (pp.47-56). New York: Palisades Institute.

- Howell, W. C., & Kraft, C. L. (1959). *Size, blur, and contrast as variables affecting the legibility of alpha-numeric symbols on radar-type displays*(Tech. Report WADC-TR-59-536). Wright Patterson Air Force Base, OH: Laboratory of Aviation Psychology.
- Hufnagel, R. E. (1965) *A search for a summary measure of image quality. Part II.* Paper presented at the Annual Meeting of the Optical Society of America, Philadelphia.
- Human Factors Society Working Group (1986). *American National Standard for Human Factors Engineering of Visual Display Terminal Workstations.* Unpublished Draft Standard.
- Information Display*, (1985,June) SID "Best Paper" displays full color LC video panel, 16-17.
- Isensee, S. H., & Bennett, C, A. (1983). The perception of flicker and glare on computer CRT displays. *Human Factors*, 25, 177-184.
- Keesee, R. L. (1976). *Prediction of modulation detectability thresholds for line-scan displays* (Tech. Report AMRL-TR-76-38). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.
- Klingberg, C. S., Elworth, C. S., & Filleau, C. R. (1970). *Image quality and detection performance of military photointerpreters*(Tech. Report D162-10323-1). Seattle: Boeing Company.
- Knowles, W. B., & Wulfeck, J. W. (1972). Visual performance with high-contrast cathode-ray tubes at high levels of ambient illumination. *Human Factors*, 14, 521-532.
- Krebs, M. J. (1978). Design principles for the use of color displays. *S.I.D International Symposium Digest of Technical Papers, IX*, 28-29.
- Laycock, J. (1985a). The legibility of passive displays. In *Proceedings of the Society for Information Display*, 26, (pp. 89-93). New York: Palisades Institute.
- Laycock, J. (1985b, April). The effect of picture element failure on the legibility of a matrix display image. *Displays*, 70-77.
- Lehmann, W. (1980, April). Electroluminescent large-area image display. *Displays*, 29-38.
- Lehrer, N. H. (1985). The challenge of the cathode-ray tube. In L. E. Tannas (Ed.), *Flat-panel displays and CRTs* (pp. 138-176). New York: Van Nostrand Reinhold Co.
- Linfoot, E. H. (1960). *Fourier methods in optical image evaluation.* London: Focal Press.
- Lippert, T. M. (1984). *Color contrast effects for a simulated CRT head-up display.* Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

- Lippert, T. M. (1985). *Unitary suprathreshold color-difference metrics of legibility for CRT raster imagery*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Lippert, T. M., Farley, W. W., Post, D. L., & Snyder, H. L. (1983). Color contrast effects on visual performance. *S.I.D. International Symposium Digest of Technical Papers, XIV*, 170-171.
- Maddox, M. E. (1977). *Prediction of information transfer from computer-generated dot-matrix displays*. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Maddox, M. E., Burnette, J. T., & Gutmann, J. C. (1977). Font comparisons for 5X7 dot-matrix characters. *Human Factors, 19*, 89-93.
- Marsetta, M., & Shurtleff, D. A. (1966). *Studies in display symbol legibility Part XIV: The legibility of military map symbols on television*. (Tech. Report MTR-264). Hanscom Field Bedford, Mass: Decision Science Laboratory.
- McCormick, E. J. (1976). *Human factors in engineering and design*(4th ed.). New York: McGraw Hill.
- McCormick, E. J., & Sanders, M. S. (1982). *Human factors in engineering and design*(5th ed.). New York: McGraw Hill.
- McTyre, J. H. (1982). Legibility comparison of 7 by 7 and 7 by 9 CRT character dot matrices. In *Proceedings of the Human Factors Society 26th Annual Meeting*(pp.710-714). Santa Monica, CA: Human Factors Society.
- Nicholson, M. M. (1984). Electrochromic flat-panel multicolor displays. *Information Displays, 2*, 4-14.
- O'Donnell, R. D., & Gomer, F. E. (1976). *Comparison of human information processing performance with dot and stroke alphanumeric characters* (Tech. Report AMRL-TR-75-95). Wright Patterson Airforce Base, OH: Aerospace Medical Division.
- Overington, I. (1975). Some considerations on the role of the eye as a component of an imaging system. *Optica Acta, 22*, 365-374.
- Overington, I. (1976). *Vision and acquisition*. London: Pentech Press.
- Overington, I. (1982). Towards a complete model of photopic visual threshold performance. *Optical Engineering, 21*, 2-13.
- Pastor, J. R., & Uphaus, J. S. (1982). Significant reading failures in 7X9 dot-matrix ASCII numbers with two percent dot loss. *S.I.D. International Symposium Digest of Technical Papers, XIII*, 198-199.

- Payne, S. J. (1983). Readability of liquid crystal displays: A response surface. *Human Factors*, 25, 185-190.
- Penz, P. A. (1985). Nonemissive displays. In L.E. Tannas (Ed.), *Flat-panel displays and CRTs*(pp. 415-457). New York: Van Nostrand Reinhold Co.
- Peters, G. L., & Barbato, G. J. (1976). *Information Processing of dot-matrix displays* (Tech. Report AFFDL-TR-76-82). Wright-Patterson Air Force Base, Ohio: Air Force Flight Dynamics Laboratory.
- Plath, D. W. (1970). The readability of segmented and conventional numerals. *Human Factors*, 12, 493-498.
- Post, D. L. (1983). *Color contrast metrics for complex images*. Unpublished doctoral dissertation Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Post, D. L. (1985). Effects of color on CRT symbol legibility. *S.I.D International Symposium Digest of Technical Papers, VVI*, 196-199.
- Post, D. L., Costanza, E. B., & Lippert, T. M. (1982). Expressions of color contrast as equivalent achromatic contrast. In *Proceedings of the Human Factors 26th Annual Meeting* (pp. 581-585). Santa Monica, CA: Human Factors Society.
- Pratt, W. K. (1978). *Digital image processing*. New York: Wiley.
- Riley, T. M., & Barbato, G. J. (1978). Dot-matrix alphanumeric viewed under discrete element degradation. *Human Factors*, 20, 473-479.
- Rosell, F. A. (1971). *Analysis of electro-optical imaging sensors* (Tech. Report ADTM 105). Baltimore: Westinghouse Electric Corporation, Systems Development Division.
- Rosell, F. A., & Willson, R. H. (1973). Recent psychophysical experiments and the display signal-to-noise ratio concept. In L. M. Biberman (Ed.), *Perception of displayed information* (pp. 167-232). New York: Plenum.
- Rupp, B. A. (1981). Visual display standards: A review of issues. In *Proceedings of the Society for Information Display*, 22, (pp. 63-72), New York: Palisades Institute.
- Schade, O. H. (1953). 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.
- Schindler, R. A. (1976). *Optical power spectrum analysis of display imagery. Phase I: Concept validity* (Tech. Report AMRL-TR-76-96). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.

- Schindler, R. A. (1979). *Optical power spectrum analysis of processed imagery* (Tech. Report AMRL-TR-79-29). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Lab.
- Sekuler, R., Tynan, P. D., & Kennedy, R. S. (1981). *Sourcebook of temporal factors affecting information transfer from visual displays* (Tech. Report 540). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.
- Semple, C. A., Heapy, R. J., Conway, E. J., & Burnett, K. T. (1971). *Analysis of human factors data for electronic flight display systems* (Tech. Report AFFDL-TR-70-174). Wright-Patterson Air Force Base, Ohio: Flight Dynamics Laboratory.
- Shannon, C. E. & Weaver, V. (1949). *The mathematical theory of communication*. Urbana, Ill.: University of Illinois Press.
- Sherr, S. (1979). *Electronic Displays* New York: John Wiley and Sons.
- Shurtleff, D. A. (1970a). *Studies of display symbol legibility XXI. The relative legibility of symbols formed from matrices of dots*. (Tech. Report ESD-TR-69-432). L.G. Hanscom Field, Bedford, MASS: Electronic Systems Division.
- Shurtleff, D. A. (1970c). *Studies of display symbol legibility XXIV. The relative legibility of special symbols formed from different matrices and the legibility of overprinted special symbols*. (Tech. Report ESD-TR-69-439). Hanscom Field, Bedford, MASS: Electronic Systems Division.
- Shurtleff, D. A. (1974). *Legibility Research*. In *Proceedings of the Society for Information Display* (pp. 41-51). New York: Palisades Institute.
- Shurtleff, D. A. (1980). *How to make displays legible*. La Mirada, CA: Human Interface Design Publisher.
- Shurtleff, D. A., Marsetta, M., & Showman, D. (1966). *Studies of display symbol legibility Part IX. The effects of resolution, size, and viewing angle of legibility* (Tech. Report ESD-TR-65-411). L. G. Hanscom Field, Bedford, Mass: Electronics System Division.
- Shurtleff, D. A., & Owen, D. (1966b). *Studies of display symbol legibility Part VII: Comparison of displays at 945- and 525- Line Resolutions* (Tech. Report ESD-TR-65-137). L. G. Hansom Field, Bedford, Mass. Electronic Systems Division.
- Smith, S. L. (1978, September). *Letter size and legibility*. Paper presented at the NATO Conference on Visual Presentation of Information, Het Vennebos, The Netherlands.
- Snyder, H. L. (1967). *Low-light-level TV viewfinder simulation program. Phase A: State-of-the-art review and simulation plans* (Tech. Report AFAL-TR-67-293). USAF.
- Snyder, H. L. (1973). *Image quality and observer performance*. In L. M. Biberman (Ed.), *Perception of displayed information* (pp. 87-118). New York: Plenum.

- Snyder, H. L. (1974). Image quality and face recognition on a television display. *Human Factors*, 16, 300-307.
- Snyder, H. L. (1976). *Visual search and image quality: Final report* (Tech. Report AMRL-TR-76-89). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.
- Snyder, H. L. (1980). *Human visual performance and flat panel display image quality* (Tech. Report HFL-80-1). Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Snyder, H. L. (1983). *Quality metrics of digitally derived imagery and their relation to interpreter performance: VIII Final Report* (Tech. Report HFL-83-1). Blacksburg, VA: Virginia Polytechnic Institute and State University.
- Snyder, H. L. (1985). Image quality: Measures and visual performance. In L. E. Tannas (Ed.), *Flat-panel displays and CRTs* (pp. 70- 90). New York: Van Nostrand.
- Snyder, H. L., Beamon, W. S., Gutmann, J. C., & Dunsker, E. D. (1980). *An evaluation of the effect of spot wobble upon observer performance with raster scan displays* (AMRL-TR-79-81). Wright-Patterson Air Force Base, OH: Air Force Aerospace Medical Research Laboratory.
- Snyder, H. L., Keesece, R. L., Beamon, W. S., & Aschenbach, J. R. (1974). *Visual search and image quality* (Tech. Report AMRL-TR-73-114). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.
- Snyder, H. L. & Maddox, M. E. (1978). *Information transfer from computer-generated, dot-matrix displays* (Tech. Report HFL-78-3). Virginia Polytechnic Institute and State University, Blacksburg, Va.
- Snyder, H. L., Shedivy, D. I., & Maddox, M. E. (1980). *Quality metrics of digitally derived imagery and their relation to interpreter performance: III. Subjective scaling of hard-copy digital imagery*(Tech. Report HFL-81-3). Blacksburg, VA: Virginia Polytechnic Institute and State University.
- Snyder, H. L. & Taylor, D. F. (1976). *Computerized analysis of eye movements during static display visual search* (Tech. Report AMRL-TR-75-91). Wright-Patterson Air Force Base, Ohio: Aerospace Medical Research Laboratory.
- Snyder, H. L., & Taylor, G. B. (1979). The sensitivity of response measures of alphanumeric legibility to variations in dot-matrix display parameters. *Human Factors*, 21, 457-471.
- Snyder, H. L., Turpin, J. A., & Maddox, M. E. (1980). *Quality metrics of digitally derived imagery and their relation to interpreter performance: II. Effects of blur and noise on hard-copy interpretability*(Tech. Report HFL 81-2). Blacksburg, VA: Virginia Polytechnic Institute and State University.
- Stein, I. H. (1980). The effect of active area on the legibility of dot-matrix displays. In *Proceedings of the Society for Information Display*, 21, (pp. 17-20). New York: Palisades Institute.

- Sutton, J., & Powers, J. (1984, April). Bringing new technology to an old industry. *Information Display*, 4-8.
- Tannas, L.E. (1985). Electroluminescent displays. In L.E. Tannas (Ed.), *Flat-panel displays and CRTs* (pp. 238-288). New York: Van Nostrand Reinhold Co.
- Task, H. L. (1979). *An evaluation and comparison of several measures of image quality for television displays* (Tech. Report AMRL-TR-79-7). Wright-Patterson Air Force Base, OH: Air Force Aerospace Medical Research Laboratory.
- Task, H. L. & Verona, R. W. (1976). *A new measure of television display quality relatable to observer performance* (Tech. Report AMRL-TR-76-73). Wright-Patterson Air Force Base, OH: Air Force Aerospace Medical Research Laboratory.
- van Meeteren, A. (1973). *Visual aspects of image intensification*. Soesterberg, The Netherlands: Institute for Perception TNO.
- Vanderkolk, R. (1976). Dot matrix Symbology. In *Human Factors in dot-matrix displays* (Tech. Report AFFDL-TR-48). Wright Patterson Air Force Base, OH: Air Force Flight Dynamics Laboratory.
- Vartebedian, A. G. (1970a). Effect of parameters of symbol formation on legibility. *Journal of the Society for Information Display*, 7, 23-26.
- Vartebedian, A. G. (1971a). Legibility of symbols on CRT displays. *Applied Ergonomics*, 2., 130-132.
- Vartebedian, A. G. (1971b). The effects of letter size, case, and generation method on CRT display search time. *Human Factors*, 13, 363-368.
- Wagner, D. W. (1977). *Color coding an annotated bibliography* (Tech. Report NWC TP 5922). China Lake, CA: Naval Weapons Center.
- Walraven, J. (1985a). Perceptual problems in display imagery. *S.I.D. International Symposium Digest of Technical Papers, XVI*, 192-195.
- Walraven, J. (1985b, January). The colours are not on the display: a survey of non-veridical perceptions that may turn up on a colour display. *Displays*, 35-41.
- Williams, R. D. (1981). Design considerations for distance viewed dot-matrix displays. In *Proceedings of the Human Factors Society 25th Annual Meeting* (PP.26-29). Santa Monica, CA: Human Factors Society.
- Winkler, R. E. (1979). *Readability of electronic displays* (Tech. Report 79-237T). Manhattan, Kansas: Kansas State University.

ANNOTATED BIBLIOGRAPHY

- Abramson, S. R., & Snyder, H. L. (1984). *Operator performance on flat-panel displays with line and cell failures* (Tech. Report HFL-83-3). Blacksburg: Virginia Polytechnic Institute and State University, Human Factors Laboratory.

This comprehensive study investigated the effects of dot and line failures on reading performance manipulating failure type, failure mode, percent failure, font, and case. Complex interactions between variables were found. Cell failures resulted in the longest reading speeds, and as the percent of failures increased response time performance increased.

- Albert, D. E. (1975). *Prediction of intelligibility of contextual dot-matrix characters*. Unpublished Master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Contextual and noncontextual (anagrams) words were investigated under varying levels of display luminance and character size. Presentations of contextual letters improve performance over noncontextual letter presentations under degraded conditions of low luminance and small character size.

- Aldersey-Williams, H. (1985, December). LCDs: a bright new face. *High Technology*, 40-42.

This article discusses LCD technology and reports the major technological breakthroughs of various companies such as Data General, Citizen Watch, Panel Vision, Suwa Seikosha, General Electric, Sanyo, and others. Market forecasts are also included.

- Almagor, M., Farley, W. W., & Snyder, H. L. (1979). *Spatiotemporal integration in the visual system* (Tech. Report AMRL-TR-78-126). Wright-Patterson Air Force Base, OH: Air Force Aerospace Medical research Laboratory.

This study shows that the integration time of the visual system is not constant and varies greatly with adapting luminance, individual observer differences, and the noise level of the display. It also is determined that the visual system trades off spatial integration with temporal integration to obtain an optimum image.

- Avery, T. E., & Berlin, G. L. (1985). *Interpretation of Aerial Photographs*. Minneapolis, MINN: Burgess, Publishing Co.

This is a textbook that discusses state-of-the-art aerial photography. A chapter on digital imagery is included.

- Banks, W. W., Gertman, D. I., & Petersen, R. J. (1982). *Human Engineering Design Considerations for Cathode Ray Tube-Generated Displays* (Tech. Report NUREG/CR-2496). Washington D. C.: U. S. Nuclear Regulatory Commission.

This report reviews twelve source documents for recommendations for 22 display variables relevant to legibility. Excerpts from *Design Handbook For Imagery Interpretation Equipment* (1975) by Richard J. Farrell and John M. Booth, eds, are also included in this document.

- Bauer, D., & Cavonius, C. R. (1980). Improving the legibility of visual display units through contrast reversal. In E. Grandjean and E. Vigliani (Eds.), *Ergonomic aspects of visual display terminals* (pp. 137-142). London: Taylor and Francis.

Two experiments were conducted in which observers were asked to identify letters on both positive and negative contrast displays. Error rate and time to perform to the task were

recorded. Results indicated that errors and time were significantly lower for the negative contrast condition.

- Beamon, W. S. & Snyder, H. L. (1975). *An experimental evaluation of the spot wobble method of suppressing raster structure visibility*(Tech. Report AMRL-TR-75-63). Wright-Patterson Air Force Base, OH: Air Force Aerospace Medical Research Laboratory.

This experiment evaluated changes in operator performance as indicated by two performance measures in a simulated air-to-ground search task. The performance parameters were evaluated at four spot wobble amplitudes. The main findings were that spot wobble had no significant effect on the number of correct responses.

- Beaton, R. J. (1984). *A human-performance based evaluation of quality metrics for hard-copy and soft-copy digital imagery systems*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

This study examined a large selection of quality metrics for hard-copy and soft-copy digital imaging systems. The results indicated that a signal-to-noise-ratio metric provided the best description of human performance across two different photointerpretation tasks. The results also indicate that metrics which incorporate the human CSF correlate higher with human performance scores than do the non-perceptually weighted metrics.

- Bescos, J. & Santamaria, J. (1977). Formation of color images. Optical transfer functions for the tristimulus values. In R. Shaw (Ed.), *SPSE Conference Proceedings* (pp. 38-47). Washington, D. C.: Society of Photographic Scientists.

Three optical transfer functions corresponding to the three tristimulus values are introduced. Their possible applications are discussed.

- Blumenthal, A. H. & Campana, S. B. (1981). An improved electro-optical image quality summary measure. In *Proceedings of the Society of Photo-Optical Instrumentation Engineers, Image Quality, 310*, (pp. 43-52). Bellingham, WA: SPIE.

This paper describes efforts to develop and validate a model which is more effective than the limiting resolution in predicting image quality. Implications in performance prediction modeling and electro-optic sensor design trade-offs between sensitivity and resolution are also discussed.

- Borough, H. C., Fallis, R. F., Warnock, R. H., & Britt, J. H. (1967). *Quantitative determination of image quality* (Tech. Report D2-114058-1). Boeing Company.

The MTF metric was related to subjective estimates of image quality obtained from a large number of trained image interpreters using photographic imagery. The MTF metric was shown to be strongly related to subjective estimates of image quality.

- Bourdy, C., Vienot, F., Monot, A., & Chiron, A. (1985, January). Spatial contrast detection for colour patterns under selective chromatic adaptation. *Displays*, 43-51.

This study investigated the shape of the contrast sensitivity function of the visual system with red, green, and blue stimuli under selective chromatic adaptation using foveal and peripheral vision. Results are compared with achromatic tests. An interpretation of the contrast sensitivity function shape changes is discussed relative to neurological properties of achromatic and chromatic mechanisms.

- Bracewell, R. (1965). *The Fourier transform and its applications*. New York: McGraw Hill.

This is a text intended for those who are concerned with applying Fourier transforms to physical situations rather than furthering the mathematical subject as such.

- Brown, J. L. (1965). Flicker and intermittent stimulation. In C. H. Graham (Ed.), *Vision and visual perception* (pp. 251- 320). New York: Wiley.

Review of research on flicker and discussion of theoretical formulations.

Buckler, A. T. (1977). *A review of the literature on the legibility of alphanumeric on electronic displays*. (TM 16-77). Aberdeen, MD: U.S. Army Human Engineering Laboratory.

A review of alphanumeric legibility research for electronic displays. Recommendations for several display parameters are offered.

Burnette, J. T. (1976). *Optimal element size-shape-spacing combinations for a 5X7 dot-matrix visual display under high and low ambient illumination*. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg VA.

Operator performance on three separate tasks was measured under various combinations of element size, shape, and spacing, and illumination. Significant effects for all variables were found as well as many interactions. A square element shape resulted in the best performance for both reading and search tasks.

Campbell, F. W., Kulikowski, J. J., & Levinson, J. (1966). The effect of orientation on the visual resolution of gratings. *Journal of Physiology*, 187, 427-436.

Experiments are described which evaluate the optical and neurophysiological factors involved in the visual resolution of sinusoidal gratings presented at different orientations.

Campbell, F. W. & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, 197, 551-566.

Results are presented which show the existence within the nervous system of linearly operating independent mechanisms which are selectively sensitive to limited ranges of spatial frequencies.

Carlson, C. R. & Cohen, R. W. (1978). *Visibility of displayed information: Image descriptors for displays* (Tech. Report ONR-CR213-120-4F). Arlington, VA: Office of Naval Research.

These researchers have developed a series of Discriminable Difference Diagrams which correspond to a variety of display conditions. The image quality metric derived from this approach is the sum of the just noticeable differences under the MTF.

Carlson, C. R., Cohen, R. W., & Gorog, I. (1977). Visual processing of simple two-dimensional sine-wave luminance gratings. *Vision Research*, 17, 351-358.

Threshold contrast sensitivity measurements were made using both one- and two-dimensional sinusoidal luminance gratings. The results are consistent with the existence of orientation-specific channels in the human visual system.

Carter, R. C. (1982). A design tool for color displays. In *Proceedings of the Human Factors Society 26th Annual Meeting*. (pp.589). Santa Monica, CA: Human Factors Society.

A one page descriptive summary of an algorithm designed to aid in selecting colors for enhanced contrast for multichromatic displays.

Chao, B. P. (1983). *Human performance evaluations of selected image enhancement/restoration techniques* Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

In this study ten enhancement/restoration processes were applied to digital images which had varying degrees of blur and noise. Professional photointerpreters subjectively rated the image quality and were also required to perform information extraction tasks. Processes were found to improve the information content of degraded images as expected. The author summarizes the usefulness of each technique for removing various amounts of blur and noise. The results of this study were compared to other studies using both hardcopy and softcopy image displays.

Charman, W. N. & Olin, A. (1965). Tutorial: Image quality criteria for aerial camera systems. *Photographic Science and Engineering*, 9, 385- 397.

The usefulness of different criteria for rating the quality of aerial camera systems is discussed in terms of the conditions under which aerial photographs are taken and interpreted. A single-number criterion of image quality, the threshold quality factor, is developed which is derived from the modulation transfer function (MTF) but which also takes some account of all the relevant photographic conditions.

Cohen, R. W. & Gorog, I. (1974). Visual capacity: An image quality descriptor for display evaluation. In *Proceedings of the Society for Information Display*, 15,(pp. 53-62). New York: Palisades Institute

The visual capacity metric is described. It is thought of as the total number of edges that can be perceived by an observer located at a given distance from the display. Specific examples are given.

Cornog, D. Y., & Rose, F. C. (1967). *Legibility of alphanumeric characters and other symbols: II. A reference handbook*. Washington: U. S. Government Printing Office, National Bureau of Standards Miscellaneous Publication 262-2.

A classic review of the legibility research dating back from the 1940's through 1967. Each report is thoroughly annotated.

Craford, M. G. (1985). Light-emitting diode displays. In L. E. Tannas (Ed.), *Flat-panel displays and CRTs* (pp. 289- 331). New York: Van Nostrand Reinhold Co.

A technical description of LED devices including history, physics and chemistry, and performance characteristics. A section is devoted to the use of LEDs as display devices.

Dahljelm, H. D. (1976). *Investigation of alphanumeric symbol legibility determination by use of Fourier spatial frequency components*. Unpublished master's thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio.

The legibility of alphanumeric sets was predicted by using the two- dimensional Fourier spatial frequency components of the symbols.

Dainty, J. C. & Shaw, R. (1974). *Image science*. New York: Academic Press.

This is a text which provides a fundamental treatment of the principles and analysis of imaging processes and the evaluation of their images.

DeCicco M. J. (1983). *An experimental determination of the effects of dot-matrix/character size and font on reading times and eye movements*. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

The investigator manipulated matrix/character size and font to determine the effects on reading speed and eye movements. Results indicated significant practice effects for three of the matrix/character sizes (7 X 9, 7 X 9 (equal in size to 5 X 7), and 9 X 11 (equal in size to 5 X 7)). Character size was also found to have some effect on the duration of eye fixations and the duration of saccades.

De Corte, W. (1985, April). High contrast sets of colours for colour CRTs under various conditons of illumination. *Displays*, 95-100.

This article discusses two methods for expanding the Carter and Carter algorithm for computing high-contrast colors on CRTs, to include the effects ambient illumination. Application examples are given.

Donohoo, D. T., & Snyder, H. L. (1985). Accommodation during color contrast. *SID International Symposium Digest of Technical Papers*, XVI, 200-203.

This study investigated accommodation responses to chromatic targets on chromatic backgrounds. Chromatic aberration was found to affect accommodation response, but the authors found that only a few target background combinations should be avoided.

- Duncan, J., & Konz, S. (1974). Effect of ambient illumination on legibility of displays of liquid-crystals and light-emitting diodes. In *Proceedings of the Human Factors Society 18th Annual Meeting* (pp.102-108). Santa Monica, CA: Human Factors Society.

Two liquid crystal displays and three light-emitting diode displays were evaluated under varying levels of illumination. Five different response measures were used; recognition time of single digits, no-error viewing distance, preferred viewing distance, preferred illumination level, and display preference.

- Ellis, B., Burrell, G. J., Wharf, J. H., & Hawkins, T. D. F. (1975). The format and colour of small matrix displays for use in high ambient illumination. In *Proceedings of the Society for Information Display, 16*, (pp. 250-258). New York: Palisades Institute.

The legibility of two sets of 5 X 7 characters each with different dot spacing ratios (1:1 and 2.5:1) were evaluated under high ambient illumination (10,000 ft-c.). Characters made of the larger dot size ratio had a lower recognition rate (7.8%), but not significantly so. The authors also evaluated the legibility of characters using red versus green display luminance under 10,000 ft-c. The dot space ratio was held constant at 1.75:1. The recognition rate for green characters was significantly longer than for red characters. Authors suggest a green display should have three times the luminance of a red display under high ambient illumination.

- Farrell, R. J., & Booth, J. M. (1975). *Design handbook for imagery interpretation equipment* (Tech. Report AD/A-025 453). Boeing Aerospace Company, Seattle, WA: U. S. Department of Commerce National Technical Information Service.

A handbook written to aid in the design and procurement of imagery interpretation equipment. Sections include optical imagery display, electro-optical imagery displays, special imagery displays, workstation design, and facilities. An excellent source document

- Flegal, R. T., Coovert, R.E., & King, C. N. (1985). A large-area electroluminescent display. *S.I.D. International Symposium Digest of Technical Papers, XIV*, 213-214.

Driving architecture, power requirements, and panel design of a 640 X 400 electroluminescent display are reported.

- Fujita, Y., Kuwata, J., Nishikawa, M., Tohda, T., Matsuoka, T., Abe, A., & Nitta, T. (1984). Large-scale ac thin film electroluminescent display panel. In *Proceedings of the Society for Information Display, 25*, (pp.177-182). New York: Palisades Institute.

The chemical structure and breakdown phenomena of a new dielectric AC thin film is reported and a TFEL panel is described.

- Gaskill, J. D. (1978). *Linear systems, Fourier transforms, and optics*. New York: Wiley.

Complete treatments of general harmonic analysis, linear systems, convolution, and Fourier transformations for one- and two-dimensional signals are presented with applications to optics.

- Giddings, B. J. (1972). Alphanumeric for raster displays. *Ergonomics, 15*, 65-72.

This study investigated the legibility of words and digits to determine an appropriate character size. The required subtended visual angle of digits was found to be significantly higher than that required for alpha characters.

- Goodman, L. A. (1974). The relative merits of LEDs and LCDs. In *Proceeding of the Society of Information Display, 16*, (pp.8-15). New York: Palisades Institute.

This article reviews the operating principles of liquid crystal displays and light emitting diode displays. The displays are compared in terms of visual appearance, power dissipation, response times, temperature dependence, circuit compatibility, and packaging.

Goodman, L. A. (1975). Passive liquid displays: liquid crystals, electrophoretics, and electrochromics. *IEEE Transactions on Consumer Electronics, CE-21*, 247-259.

This article reviews the principle operations of liquid crystal, electrophoretic, and electrochromic displays. Each display is compared in terms of visual appearance, power dissipation, reliability, response times, circuit compatibility, and temperature dependence.

Gould, J. D. (1968). Visual factors in the design of computer-controlled CRT displays. *Human Factors, 10*, 359-376.

This article reviews several display parameters including luminance, contrast, chromaticity, resolution, and alphanumeric character design for CRT displays. Recommendations are given. Although written 18 years ago this article is still very relevant.

Graff, G. (1985, December). Electroluminescence: better picture, higher cost. *High Technology, 42-45*.

Electroluminescent display technology is described in layman terms, and currently available EL displays are described. Leading techniques for multicolor EL displays are briefly discussed.

Grandjean, E., & Vigliani, E. (Eds.). (1980). *Ergonomic aspects of visual display terminals*. London: Taylor and Francis.

This book is a proceedings of an international workshop held in Milan in March 1980. Papers included are on the following topic areas; physical characteristics of VDTs, visual functions, visual impairments, performance at VDTs, postural problems, psychosocial aspects, practical experiences, and ergonomic design and guidelines.

Granger, E. M. (1974, May). Subjective assessment and specification of color image quality. In *Proceedings of the Society of Photo-Optical Instrumentation Engineers, 46*, (pp. 86-91). Bellingham, WA: SPIE.

The subjective quality factor metric is presented. Results are presented which test the proposed metric using color and black-and-white photographs and subjective image quality.

Granger, E. M. & Cupery, K. N. (1972). An optical merit function (SQF), which correlates with subjective image judgements. *Photographic Science and Engineering, 16*, 221- 230.

The SQF metric is studied for a wide variety of MTF shapes. It was shown to correlate highly with the subjective rank of the quality of photographs.

Granrath, D. J. (1981). The role of human visual models in image processing. In *Proceedings of the IEEE, 69*, (pp. 552-561).

Mathematical models of the mechanisms by which the human eye forms a neural image of the outside world for transmission along the optic nerve are presented. Their usefulness and limitations are discussed.

Gutmann, J. C., Snyder, H. L., Farley, W. W., & Evans, J. E. (1979). *An experimental determination of the effect of image quality on eye movements and search for static and dynamic targets* (Tech. Report AMRL-TR-79-51). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.

Two experiments investigated the effects of the quality of a televised image on eye movements and search-related dependent measures. The MTF metric was correlated with observer performance.

Hairabedian, B., Lorenzen, J., Perry, C., Pleshko, P., Rita, R., & Sherk, T. (1985). Large screen AC plasma display technology overview. *Display and Imaging Technology*, 1, 5-24.

IBM's 581 AC Plasma Display Technology is described. This technology was developed to provide large screen, multiple image-format capability

Hall, C. F. (1981, August). Subjective evaluation of a perceptual quality metric. In *Proceedings of the Society for Photographic Instrumentation Engineers, Image Quality*, 310 (pp. 200-204). Bellingham, WA: SPIE.

The PMSE metric is shown to correlate highly with the subjective rank ordering of a black-and-white image.

Hemingway, J. C., & Erickson, R. A. (1969). Relative effects of raster scan lines and image subtense on symbol legibility on television. *Human Factors*, 11, 331-338.

In this study symbol size and the number of raster lines per symbol height were manipulated to determine their effects upon observer performance for identifying geometric symbols. Authors conclude that at least 8 raster lines with a symbol size of 10 min of arc was required for good symbol legibility. An equation for determining symbol size and number of raster lines is presented.

Howard, W. E. (1981). Electroluminescent display technologies and their characteristics. In *Proceedings of the Society for Information Display*, 22, (pp.47-56). New York: Palisades Institute.

The operating characteristics of the four electroluminescent technologies are described.

Howell, W. C., & Kraft, C. L. (1959). *Size, blur, and contrast as variables affecting the legibility of alpha-numeric symbols on radar-type displays*(Tech Report WADC-TR-59-536). Wright Patterson Air Force Base, OH: Laboratory of Aviation Psychology.

This study is a classic experiment quoted often in the literature. Mackworth font characters and numerals were presented using a film strip. Four levels each of character size, blur, and contrast were manipulated. Results are discussed in terms of an information metric expressed in entropy units, accuracy, and processing time. Results indicated that when characters are larger than 16 min of arc, contrast and blur has little effect except at very low levels of these variables.

Huddleston, H. F. (1970). An initial examination of 5x7 matrix alphanumerics for digital television. *IAM Report No. 496*.

This study describes three experiments conducted to determine the best alphanumeric design for a 5 x 7 matrix to be displayed on digital television. An alphanumeric font was recommended.

Huddleston, H. F. (1974). A comparison of two 7x9 matrix alphanumeric designs for T.V. displays. *Applied Ergonomics*, 5.2, 81-83.

This study compared an RAE font designed by the author and a font designed by Vartebedian (1970) for use in airborne displays. The author concluded that the RAE font provided better performance in terms of recognition errors than Vartebedian's font. The results do not necessarily generalize to discrete or dot-matrix alphanumerics.

Hufnagel, R. E. (1965) *A search for a summary measure of image quality. Part II.* Paper presented at the Annual Meeting of the Optical Society of America, Philadelphia.

Hufnagel describes his Q3 metric.

Human Factors Society Working Group (1986). *American National Standard for Human Factors Engineering of Visual Display Terminal Workstations.* Unpublished Draft Standard.

Technical standard of specifications for visual display terminals. This document discusses the working environment, visual display design, workstation design, keyboard design, and display measurement techniques.

Information Display, (1985, June) SID "Best Paper" displays full color LC video panel, 16-17.

This article discusses construction and specifications of a 4.25-in diagonal LC panel. The article is based on a paper presented at the *Society for Information Display International Symposium* in 1984.

Isensee, S. H., & Bennett, C. A. (1983). The perception of flicker and glare on computer CRT displays. *Human Factors*, 25, 177-184.

In this experiment ambient illumination, display luminance, and display polarity were manipulated. The off-axis angle at which subjects first noticed flicker was recorded, as was subjects discomfort ratings. Display luminance had the greatest impact on discomfort ratings. Higher display luminances resulted in smaller angles for detecting flicker. Lower illumination levels also resulted in smaller angles. Methods for reducing flicker and glare are discussed.

Kaneko, E., Kawakami, H., Hanmura, H., & Saito, S. (1982). A pocket-size Liquid-Crystal TV Display. In *Proceedings of the Society for Information Display*, 23, (pp. 3-8). New York: Palisades Institute.

Construction and operating characteristics of a LC display with an effective display area of 45 x 60 mm (3-in diagonal) and 120 X 160 pixels.

Kasahara, K., Sakai, K., Komatsubara, Y., Saito, A., Ide, K., Matsumoto, S., & Hori, H. (1983). A 480 X 480 element dichroic dye MOS-LCD. *S.I.D. International Symposium Digest of Technical Papers, XIV*, 150-151. 150-151.

Configuration, fabrication and operation of a dynamic-scattering liquid-crystal T.V. using MOS array is described.

Kasahara, K., Yanagisawa, T., Sakai, K., Adachi, T., Inoue, K., Tustsumi, T., & Hori, H. (1981). A liquid-crystal display panel using an MOS array with gate-bus drivers. In *Proceedings of the Society for Information Display*, 22, (pp. 318-322). New York: Palisades Institute.

Technical description of a 220 X 240 element liquid crystal television display using active matrix addressing. Display specifications are given.

Keesee, R. L. (1976). *Prediction of modulation detectability thresholds for line-scan displays* (Tech. Report AMRL-TR-76-38). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.

Modulation detectability threshold functions were determined for a range of system parameters typical of medium- to high-resolution television systems. Models were developed which can be used to predict threshold detectability under various display conditions.

Kelley, D. H. (1975). Spatial frequency selectivity in the retina. *Vision Research*, 15, 665-672.

Pre-processing of visual information is shown to occur at the retinal level.

Ketchel, J. M., & Jenney, L. L. (1968). *Electronic and optically generated aircraft displays: A study of standardization requirements*(Tech. Report 680505). Washington D.C.: JANAIR Office of Naval Research.

A literature review and design specification guide for electronic and optically generated displays. This report includes sections regarding display characteristics, symbology (eg. coding), information requirements, and a brief description of different display categories, such as direct and projected vertical situation displays, and horizontal situation displays.

Kilcoyne, K. (1983, June). New EL display technology. *Information Display*, 4-8.

This article discusses advances in EL technology and the advantages of EL displays.

Klingberg, C. S., Elworth, C. S., & Filleau, C. R. (1970). *Image quality and detection performance of military photointerpreters*(Tech. Report D162-10323-1). Seattle: Boeing Company.

The relationship between objectively measured information-extraction performance and MTF values was examined using photographic imagery.

Knowles, W. B., & Wulfeck, J. W. (1972). Visual performance with high-contrast cathode-ray tubes at high levels of ambient illumination. *Human Factors*, 14, 521-532.

In this experiment four CRTs (one standard and three high contrast) were evaluated under four levels of ambient illumination (100, 1,000, 5,000, and 10,000 ft-c), two angles of incidence (30- and 60-deg), and two angles of regard (90 and -45 deg). A gap detection task was used and working preferences were also recorded. Results indicated that none of the CRTs "washed-out", even under the 10,000 ft-c condition.

Krebs, M. J. (1978). Design principles for the use of color displays. *S.I.D International Symposium Digest of Technical Papers, IX*, 28-29.

This article provides general guidelines for using color coding on information displays.

Laycock, J. (1985a). The legibility of passive displays. In *Proceedings of the Society for Information Display*, 26, (pp. 89-93). New York: Palisades Institute.

The optical properties of various passive or non-emitting displays are briefly reviewed. Guidelines for alphanumeric legibility are also presented.

Laycock, J. (1985b, April). The effect of picture element failure on the legibility of a matrix display image. *Displays*, 70-77.

Laycock discusses the effects of dot and line failures on matrix-addressed displays. Dot and line failures were incorporated into various display pictures using an image processor. The author provides a subjective assessment of the effects of the failures on display legibility.

Lehmann, W. (1980, April). Electroluminescent large-area image display. *Displays*, 29-38.

This article discusses the operating characteristics of the four types of EL display screens, ac and dc thin film, and ac and dc powder films.

Lehrer, N. H. (1985). The challenge of the cathode-ray tube. In L. E. Tannas (Ed.), *Flat-panel displays and CRTs* (pp. 138-176). New York: Van Nostrand Reinhold Co.

CRT design, operation, and performance is covered.

Linfoot, E. H. (1960). *Fourier methods in optical image evaluation*. London: Focal Press.

Among other things, the SIR metric is described.

Lippert, T. M. (1984). *Color contrast effects for a simulated CRT head-up display*. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

This study investigated reading performance for chromatic targets on an chromatic backgrounds. Color contrast metrics were computed using different color spaces to determine the most appropriate metric.

Lippert, T. M. (1985). *Unitary suprathreshold color-difference metrics of legibility for CRT raster imagery*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Lippert, T. M., Farley, W. W., Post, D. L., & Snyder, H. L. (1983). Color contrast effects on visual performance. *S.I.D. International Symposium Digest of Technical Papers, XIV*, 170-171.

This study evaluated the 1976 C.I.E. $L^* u^* v^*$, $L^* a^* b^*$ and Cohen and Friden's W, a, b color spaces as a basis for performance metrics of color contrast. Numeral strings in Huddleston font were presented against background colors and response speed data were collected. Distance in each color space and numeral length were regressed on response speed. A simplified color space $Y, u' v'$ model resulted in the best fit for the data.

Maddox, M. E. (1977). *Prediction of information transfer from computer-generated dot-matrix displays*. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Predictive metrics of information transfer from dot-matrix displays were derived for two experiments. Dot size, shape, and spacing and ambient illumination were the variables manipulated in the first study. Fourier analysis of microphotometric scans of each condition were used as predictor variables. Multiple regression techniques were used to derive equations. The equations were sensitive to the predictor variables. An external validation study was conducted.

Maddox, M. E. (1980). Two-dimensional spatial frequency content and confusions among dot-matrix characters. In *Proceedings of the Society for Information Display, 21*, (pp.31-40). New York: Palisades Institute.

This study was divided into two phases. In the first phase four dot-matrix fonts and five matrix/character sizes were presented to observers. Recognition accuracy data were collected and character confusion matrices were generated. In phase two, characters of two of the fonts were photometrically scanned and were decomposed into component frequencies. The characters were then analyzed for similarity of spatial frequency content (2-D Fourier transform) & subsequently correlated with human performance. The similarity between any two characters was also evaluated using the Phi Correlation Coefficient which was then correlated with human performance. The 2-D Fourier transform was not found to correlated well with human performance, while the Phi Coefficient provided "moderate" correlations. The utility of the techniques for designing dot-matrix characters are discussed.

Maddox, M. E., Burnette, J. T., & Gutmann, J. C. (1977). Font comparisons for 5X7 dot-matrix characters. *Human Factors, 19*, 89-93.

This study investigated the legibility of three different dot-matrix fonts; the Lincoln/Mitre font and two fonts developed for the study (Maximum Dot and Maximum Angle). The font using the maximum number of dots possible resulted in fewer recognition errors. Results were attributed to the greater apparent contrast for the Maximum dot font, even though the contrast was the same for all fonts.

Marsetta, M., & Shurtleff, D. A. (1966). *Studies in display symbol legibility Part XIV: The legibility of military map symbols on television*. (Tech. Report MTR-264). Hanscom Field Bedford, Mass: Decision Science Laboratory.

One of the few studies which investigates symbol legibility. This study investigated the number of raster lines required to recognize map symbols. Authors found that at least 17 lines per symbol height were required. With practice, 11 lines per symbol height was satisfactory. Many technical problems occurred during the experiment; therefore, results should be used cautiously.

McCormick, E. J. (1976). *Human factors in engineering and design*(4th ed.). New York: McGraw Hill.

See McCormick and Sanders (1982).

McCormick, E. J., & Sanders, M. S. (1982). *Human factors in engineering and design*(5th ed.). New York: McGraw Hill.

A survey text of human factors. Topics include human input and mediation processes, workspace design, and the affects of the environment on human performance.

McLean, M. V. (1965). Brightness contrast, color contrast, and legibility. *Human Factors*, 7, 521-526.

This study investigated the effects of color contrast versus brightness contrast, direction of contrast, and contrast level. Colored stimuli were brightness matched with achromatic stimuli. Observers were asked to read a dial presented tachistoscopically. Results indicated that low contrast levels resulted in longer reading speeds. An interaction between type of contrast (color versus brightness) and direction of contrast was also found. Shorter reading speeds were found for the color contrast light on dark condition. Direction of contrast was not significantly different for the brightness contrast condition.

McTyre, J. H. (1982). Legibility comparison of 7 by 7 and 7 by 9 CRT character dot matrices. In *Proceedings of the Human Factors Society 26th Annual Meeting*(pp.710-714). Santa Monica, CA: Human Factors Society.

7 x 7 and 7 x 9 CRT dot-matrix characters were compared for legibility. A recognition task was employed and response time and accuracy data were collected. Confusion errors were also analyzed. There were no statistically significant differences between the two character sets. It should be noted that dot size, and character height and width were different between character sets; however, the researchers stated that the purpose of the study was to determine if a smaller character set would degrade performance. Results do not necessarily generalize to reading tasks.

McTyre, J. H., & Frommer, W. D. (1985). Effects of character/background color combinations on CRT character legibility. In *Proceedings of the Human Factors Society 29th Annual Meeting*(PP.779-781). Santa Monica, CA: Human Factors Society.

This study investigated the legibility of alphanumerics using four different target/background color combinations. The authors report that some combinations resulted in faster response times and fewer errors. Post-hoc data were not reported; therefore, it is not possible for the reader to determine which combinations are best. Color combinations varied in luminance contrast from 10:1 to 3:1 confounding contrast with the color combinations. Also, the colors were only specified by their subjective color labels.

Morozumi, S., Oguchi, K., Misawa, T., Araki, R., & Ohshima, H. (1984). 4.25-in and 1.15-in B/W and full-color LC video displays addressed by Poly-Si TFTs. *S.I.D. International Symposium Digest of Technical Papers*, XV, 316-319.

This article reports specifications of two matrix-addressed LCDs. One LCD is a 480 X 480 TFT array, one of the largest matrix-addressed LCDs to date.

Murch, G. M. (no date) Using color effectively: Designing to human specifications. Tektronix, Inc.

This article discusses the human visual system in terms of the capability to process color. Color perception is also discussed, and some guidelines for using color effectively are presented.

- Murch, G. M. (1982). Visual accommodation and convergence to multichromatic information displays. *S.I.D. International Symposium Digest of Technical Papers, XIII*, 192-193.

Murch measured visual accommodation and convergence to multichromatic colors using different CRT phosphors. The results indicated that accommodation and convergence differences were not as great as was previously found for monochrome colors. Pure unmixed phosphor colors showed a relationship between color and accommodation and convergence, while multiple (mixed) phosphors did not.

- Murch, G. M., Cranford, M., & McManus, P. (1983). Brightness and color contrast of information displays. *S.I.D. International Symposium Digest of Technical Papers, XIV*, 168-169.

This article discusses the difference between heterochromatic matching and flicker photometric matches for establishing display brightness and color contrast. An experiment was conducted using both methods to obtain perceived brightness measures. The authors argue that the heterochromatic matching procedure involves both brightness and the color component, both of which yield to the overall perception of contrast.

- Nicholson, M. M. (1984). Electrochromic flat-panel multicolor displays. *Information Displays, 2*, 4-14.

This article discusses the characteristics of electrochromic displays and applications. Electrochromic mechanisms, electrical parameters, color, and matrix addressing are also briefly discussed.

- Niina, T., Kuroda, S., Yamaguchi, T., Yonei, H., Tomida, Y., & Yagi, K. (1982). A multicolor GaP LED flat-panel display device for colorful display of letters and figures. In *Proceedings of the Society for Information Display, 23*, (pp 73-76). New York: Palisades Institute.

The structure, fabrication and various properties of a multi-colored GaP LED flat-panel display are reported.

- North, R. A., & Williges, R. C. (1971). *Video cartographic image interpretability assessed by response surface methodology* (Tech. Report ARL-71-22/AFOSR-71-8). Urbana-Champaign, IL: Aviation Research Laboratory.

Multiple regression prediction equations were obtained using Response Surface Methodology (RSM) to predict performance on a video cartographic symbol search task. Prediction equations were developed for both black and white and color TV monitors. The variables investigated were focus, display density of non-target symbols, visual angle, and lines per mm of area displayed. Performance for the color monitor was found to be a function of all four variables, while performance on black and white monitors was a function of focus and density.

- O'Donnell, R. D., & Gomer, F. E. (1976). *Comparison of human information processing performance with dot and stroke alphanumeric characters* (Tech. Report AMRL-TR-75-95). Wright Patterson Airforce Base, OH: Aerospace Medical Division.

This study compares observers performance on Sternberg's Memory Task using dot and stroke character sets. Response time and Visually Evoked Response (VER) measures were used. Results indicated no differences in information processing of dot and stroke characters. It is possible that the response measures used were not sensitive enough for the task.

- Osaka, N. (1985, July). The effect of VDU colour on visual fatigue in the fovea and periphery of the visual field. *Displays, 138-140*.

In this study visual fatigue was measured using the critical flicker frequency paradigm. VDU color and eccentricity were varied and the author reports that blue and red "strongly caused" visual fatigue as compared to green and yellow.

Overington, I. (1975). Some considerations on the role of the eye as a component of an imaging system. *Optica Acta*, 22, 365-374.

The performance of the visual system is described in terms of a series of quality functions. These quality functions are related to the quality factors of optical components external to the eye. Some of the more popular MTF-based figures of merit are considered in terms of the combined performance of an optical system and the eye.

Overington, I. (1976). *Vision and acquisition*. London: Pentech Press.

Visual acquisition is discussed in terms of photometry, image evaluation, visual optics, physiology, neurology, and psychology.

Overington, I. (1982). Towards a complete model of photopic visual threshold performance. *Optical Engineering*, 21, 2-13.

A conceptual model of photopic visual threshold performance is developed. The models which are developed are discussed in terms of the effects of image quality on visual performance. Overington discusses his visual efficiency metric.

Pastor, J. R., & Uphaus, J. S. (1982). Significant reading failures in 7X9 dot-matrix ASCII numbers with two percent dot loss. *S.I.D. International Symposium Digest of Technical Papers, XIII*, 198-199.

This study investigated the effects of percent dot loss on reading errors of ASCII numerals. Subjects viewed the stimuli for 200 ms. Results indicated a linear relationship between dot loss and reading errors.

Payne, S. J. (1983). Readability of liquid crystal displays: A response surface. *Human Factors*, 25, 185-190.

This study investigated the effects of viewing angle, level of backlight, character size, and ambient illumination on the readability of four-digit seven segment numbers presented on a reflective liquid crystal display. Error percentages were recorded and evaluated using a response surface methodology. Multiple regression prediction equations are presented. Results indicated that backlight was a strong predictor variable, adversely affecting performance as backlight increased. Character size was also found to be a strong predictor of performance, while viewing angle and ambient illumination were not.

Penz, P. A. (1985). Nonemissive displays. In L.E. Tannas (Ed.), *Flat-panel displays and CRTs*(pp. 415-457). New York: Van Nostrand Reinhold Co.

In this chapter, Penz briefly discusses the general characteristics of all types of nonemissive displays, and then discusses each type separately. Most of the chapter is devoted to description of liquid crystal devices, including the underlying physics of LCDs. Electrochromic displays (ECDs) colloidal (i.e., electrophoretic), electroactive solids, and electromechanical displays are all briefly introduced.

Peters, G. L., & Barbato, G. J. (1976). *Information Processing of dot-matrix displays* (Tech. Report AFFDL-TR-76-82). Wright-Patterson Air Force Base, Ohio: Air Force Flight Dynamics Laboratory.

This report discusses three experiments conducted to measure human cognitive processing differences using dot-matrix versus stroke characters. The authors conclude that there are small differences in processing between dot-matrix and stroke characters and that the differences are "concentrated in memorial operations".

Plath, D. W. (1970). The readability of segmented and conventional numerals. *Human Factors*, 12, 493-498.

In this study five-digit AMEL numerals, slanted segmented numerals and vertical segmented numerals were presented tachistoscopically to observers who were asked to record the numbers. Accuracy data were evaluated. Results indicated that the AMEL numerals resulted in better performance than either of the segmented numerals. There were no difference between the segmented numerals. It should be noted that strokewidth of the AMEL numerals was twice that of the segmented numerals which may have affected performance.

Post, D. L. (1983). *Color contrast metrics for complex images*. Unpublished doctoral dissertation Virginia Polytechnic Institute and State University, Blacksburg, VA.

Regression models were developed to predict response speed from color contrast for reading dot-matrix numerals presented against digitized full-color backgrounds. The color contrast difference between targets and backgrounds in three color spaces, L^*u^*v , L^*a^*b , and Wab , were determined and regressed on response speed.

Post, D. L. (1985). Effects of color on CRT symbol legibility. *S.I.D International Symposium Digest of Technical Papers*, VVI, 196-199.

The investigator was interested in determining the angular subtenses required for various symbols that differed in color and luminance. Symbols were presented on a black background. Subjects were required to perform three different tasks; symbol naming, hue naming, and a comfort legibility task. The angular subtenses required for each task are presented.

Post, D. L., Costanza, E. B., & Lippert, T. M. (1982). Expressions of color contrast as equivalent achromatic contrast. In *Proceedings of the Human Factors 26th Annual Meeting* (pp. 581-585). Santa Monica, CA: Human Factors Society.

Several experiments were conducted to compare the relationship between color contrast represented in three uniform color spaces, and achromatic contrast. Color differences in each color space were regressed on achromatic contrast settings which were obtained by having subjects adjust contrast of an achromatic pair to match the color contrast of a chromatic pair of stimuli. Results indicated that the two C.I.E. color spaces Lu^*v^* and La^*b^* are not uniform, but can be rescaled and used for specifying color contrast.

Pratt, W. K. (1978). *Digital image processing*. New York: Wiley.

This is a text for a graduate course in digital image processing. Topics covered include the mathematical representation of continuous and discrete images along with a discussion of image quality measures.

Refioglu, H. I. (Ed.). (1983). *Electronic displays*. New York: IEEE Press.

The purpose of this book is to bring together a selection of technical articles published on the subject of electronic displays. The most common and important display devices are covered.

Reingold, I. (1974). Display devices: A perspective on status and availability. In *Proceeding of the Society for Information Display*, 15, (PP. 63-73). New York: Palisades Institute.

This article reviews the different display technologies, discusses advantages and limitations, and briefly discusses future projections.

Riley, T. M., & Barbato, G. J. (1978). Dot-matrix alphanumerics viewed under discrete element degradation. *Human Factors*, 20, 473-479.

The legibility of 5 x 7 dot-matrix alphanumeric fonts was evaluated by asking subjects to remove or add dots to create character degradations. Importance values for each dot in a character were calculated. Experiments were conducted to determine the effect of degradation by removing and adding dots with both high and low importance values. No differences between fonts were found under this element degradation.

Rogers, S. P., & Gutmann, J. C. (1983). CRT symbol subtense requirements. *S.I.D. International Symposium Digest of Technical Papers, XIV*, 166-167.

Two experiments are reported that were conducted to determine subtense requirements for CRT symbols. The first experiment manipulated contrast, symbol luminance, order of luminance levels, and trials. Results indicated a significant effect of contrast, with the 2:1 level requiring greater subtended visual angles than either the 4:1 or 8:1 levels. The second experiment varied contrast, hue, order of colors, and trials holding symbol luminance constant. Contrast was the only significant effect, with the 2:1 level requiring greater subtended visual angles.

Rosell, F. A. (1971). *Analysis of electro-optical imaging sensors* (Tech. Report ADTM 105). Baltimore: Westinghouse Electric Corporation, Systems Development Division.

The performance of the unaided eye, the eye aided by simple optical aids, and the eye aided by auxiliary sensors is studied. The discussion concentrates on the thresholds of perception.

Rosell, F. A., & Willson, R. H. (1973). Recent psychophysical experiments and the display signal-to-noise ratio concept. In L. M. Biberman (Ed.), *Perception of displayed information* (pp. 167-232). New York: Plenum.

Experiments are discussed which demonstrate the prediction of the signal-to-noise ratio required in a given video bandwidth to permit various visual tasks to be conducted from displayed imagery with various levels of confidence.

Rupp, B. A. (1981). Visual display standards: A review of issues. In *Proceedings of the Society for Information Display, 22*, (pp. 63-72), New York: Palisades Institute.

This paper lists several of the recommendations put forth by various international sources for the design of visual displays. The author comments on each of the recommendations in terms of the research supporting the recommendations and their validity.

Sadacca, R., Martinek, H., & Schwartz, A. I. (1962). *Image interpretation Task--Status report, 30 June 1962* (Tech. Report 1129). Washington D. C.: US Army Personnel Research Office.

This report is a review of some of the literature between 1958 and 1962 on image interpretation. The purpose of the report was to define research problems and long range needs.

Schade, O. H. (1953). 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.

Schade develops the concepts of equivalent passbands and signal-to-deviation ratios as applied to television images.

Scheffer, T. J., Nehring, J., Kaufman, M., Amstutz, H., Heimgartner, D., & Eglin, P. (1985). 24 X 80 character LCD panel using the supertwisted birefringence effect. *S.I.D. International Symposium Digest of Technical Papers, XVI*, 120-123.

This article describes a 24 X 80 character birefringence effect (SBE) matrix display.

Schindler, R. A. (1976). *Optical power spectrum analysis of display imagery. Phase I: Concept validity* (Tech. Report AMRL-TR-76-96). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.

A basic approach to the determination of display information capacity using optical power spectrum measurements is examined mathematically and experimentally. Potential problems for practical application are identified.

Schindler, R. A. (1979). *Optical power spectrum analysis of processed imagery* (Tech. Report AMRL-TR-79-29). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Lab.

Information density measures are determined for unprocessed and processed imagery and compared with observer target recognition performance. The report describes modifications to the information density measure to improve the relationship with performance and to select the most effective processing technique.

Schlam, E. (Ed.). (January, 1983). *Advances in display technology-- Proceedings of the SPIE, 386*. Bellingham, WA: The International Society for Optical Engineering.

The papers in this book comprise the proceedings of the SPIE meeting. Topics include human factors in visible displays and image quality related factors.

Schlam, E. (Ed.). (January, 1984). *Advances in display technology IV-- Proceedings of the SPIE, 457*. Bellingham, WA: The International Society for Optical Engineering.

The papers in this book comprise the proceedings of the SPIE meeting. Flat-panel displays are described as well as user-related issues such as display perception.

Schlam, E. (1985, December). Flat-panel displays poised to displace some CRT applications. *Information Display, 11*.

A brief discussion of flat-panel constraints, applications and future prospects.

Sekuler, R., Tynan, P. D., & Kennedy, R. S. (1981). *Sourcebook of temporal factors affecting information transfer from visual displays* (Tech. Report 540). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.

A comprehensive literature review on temporal factors in vision.

Seiple, C. A., Heapy, R. J., Conway, E. J., & Burnett, K. T. (1971). *Analysis of human factors data for electronic flight display systems* (Tech. Report AFFDL-TR-70-174). Wright-Patterson Air Force Base, Ohio: Flight Dynamics Laboratory.

This report is a comprehensive literature review of over 1,000 articles relating to human factors considerations in electronic flight display systems. The articles are relevant to other uses of electronic displays as well. The following sections are included; relationship of design considerations, display size for flight control, information coding, alphanumeric design considerations, scale legibility considerations, factors affecting visual acuity, display system resolution considerations, flicker factors, legibility contrast requirements, and environmental variables.

Shannon, C. E. & Weaver, V. (1949). *The mathematical theory of communication*. Urbana, Ill.: University of Illinois Press.

The concept of information theory is developed. This forms the basis of the information content metric.

Sherr, S. (1979). *Electronic Displays* New York: John Wiley and Sons.

This book provides a description of many visual displays including CRTs, matrix-addressed flat-panel displays, and alphanumeric displays. Principles of operation are discussed. A chapter is also devoted to human perceptual factors.

Sherr, S. (1982). *Video and digital electronic displays: A user's guide*. New York: John Wiley and Sons.

The operation of video and digital displays are described for people who lack the technical background to fully understand the principles of operation.

Showman, D. J. (1966). *Studies of display symbol legibility Part X. The relative legibility of Leroy and Lincoln/MITRE alphanumeric symbols*(Tech. Report ESD-TR-66-115). L. G. Hanscom Field, Bedford, Mass: Electronic Systems Division.

Leroy and Lincoln/MITRE symbols were presented tachistoscopically to evaluate the legibility of the two fonts. Recognition accuracy data were evaluated and the Lincoln/MITRE font was found to be more legible.

Shurtleff, D. A. (1970a). *Studies of display symbol legibility XXI. The relative legibility of symbols formed from matrices of dots*. (Tech. Report ESD-TR-69-432). L.G. Hanscom Field, Bedford, MASS: Electronic Systems Division.

This study investigated the effects of matrix size on symbol legibility using the Lincoln/Mitre font. Degraded and undegraded conditions and practice effects were evaluated. In all cases but one, the larger matrix sizes (7 x 11 and 9 x 15) did not result in improved performance over the 5 x 7 matrix. These results were unexpected and the author tries to explain possible reasons for the findings.

Shurtleff, D. A. (1970b) *Studies of display symbol legibility: XXII The relative legibility of four symbol sets made with a five by seven dot-matrix*(Tech. Report ESD-TR-70-26). L. G. Hanscom Field, Bedford, Mass: Electronic Systems Division.

Four 5 x 7 dot-matrix fonts, Lincoln/MITRE, IBM 029, modified Hazeltine, and Diamond Ordinance Fuse Laboratory font, were compared under optimal and degraded viewing conditions. No one symbol set was found to be superior to the other sets.

Shurtleff, D. A. (1970c). *Studies of display symbol legibility XXIV. The relative legibility of special symbols formed from different matrices and the legibility of overprinted special symbols*. (Tech. Report ESD-TR-69-439). Hanscom Field, Bedford, MASS: Electronic Systems Division.

The legibility of 30 special symbols presented in four dot-matrix configurations were investigated. Symbols were degraded by overprinting. Rate of correct identifications and percentage of errors data were recorded. Larger matrices resulted in the best performance. Overprinting causes degraded performance.

Shurtleff, D. A. (1974). Legibility Research. In *Proceedings of the Society for Information Display* (pp. 41-51). New York: Palisades Institute.

This report discusses several parameters of legibility including number of raster lines per symbol height, dot matrix construction, circular versus elongated elements, and stroke matrix construction. Several studies by Shurtleff are reviewed.

Shurtleff, D. A. (1980). *How to make displays legible*. La Mirada, CA: Human Interface Design Publisher.

This book reviews most of the research on legibility by Shurtleff before 1980 as well as other research. Design recommendations are given.

Shurtleff, D. A., Marsetta, M., & Showman, D. (1966). *Studies of display symbol legibility Part IX. The effects of resolution, size, and viewing angle of legibility* (Tech. Report ESD-TR-65-411). L. G. Hanscom Field, Bedford, Mass: Electronics System Division.

This report investigated the visual size and the number of scan lines required for identification of the standard and a revised Leroy font. A large visual size was required for

symbols made up of 6 scan lines. For symbols made from 10 and 8 lines, visual size was approximately the same.

Shurtleff, D. A., & Owen, D. (1966a). *Studies of display symbol legibility Part VI. Leroy and Courtney symbols.* (Tech. Report ESD-TR-65-136). L. G. Hanscom Field, Bedford, Mass: Electronic Systems Division.

This study compared Courtney and Leroy alphanumeric symbols at resolutions of 12, 10, 8, and 6 scan lines per symbol height. No practical differences were found between the two fonts. A resolution of 10 lines per symbol height was recommended.

Shurtleff, D. A., & Owen, D. (1966b). *Studies of display symbol legibility Part VII: Comparison of displays at 945- and 525- Line Resolutions* (Tech. Report ESD-TR-65-137). L. G. Hanscom Field, Bedford, Mass. Electronic Systems Division.

Leroy alphanumeric characters were presented on a 945 line T.V. system at 6, 8, 10, and 12 scan lines per symbol height. Results were compared to a previous study which used a 525 line system.

Smith, S. L. (1978, September). *Letter size and legibility.* Paper presented at the NATO Conference on Visual Presentation of Information, Het Vennebos, The Netherlands.

This report summarizes field research investigating angular subtense requirements for printed material. Results indicated that the current recommendations given by various sources are valid.

Snyder, H. L. (1967). *Low-light-level TV viewfinder simulation program. Phase A: State-of-the-art review and simulation plans* (Tech. Report AFAL-TR-67-293). USAF.

Reviews the results of experiments which have been conducted to relate one or more characteristics of the visual display to the performance of the human observer in obtaining information from a search-type display.

Snyder, H. L. (1973). Image quality and observer performance. In L. M. Biberman (Ed.), *Perception of displayed information* (pp. 87-118). New York: Plenum.

Relates photometric measures of display quality to observer performance with specific emphasis on the MTF metric.

Snyder, H. L. (1974). Image quality and face recognition on a television display. *Human Factors*, 16, 300-307.

The MTF image quality metric was shown to correlate highly with measures of observer performance using a television display.

Snyder, H. L. (1976). *Visual search and image quality: Final report* (Tech. Report AMRL-TR-76-89). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.

This report presents the results of an air-to-ground television target acquisition experiment. The MTF metric was found to correlate moderately with target acquisition performance.

Snyder, H. L. (1980). *Human visual performance and flat panel display image quality* (Tech. Report HFL-80-1). Virginia Polytechnic Institute and State University, Blacksburg, VA.

This report is a survey of the pertinent visual performance, display system capability, and human engineering design requirements for flat-panel displays, as applied to U.S. Navy airborne, shipborne, and land-based systems. Current models of image quality which relate human performance to display characteristics are also discussed. Data gaps and needs are summarized.

Snyder, H. L. (1983). *Quality metrics of digitally derived imagery and their relation to interpreter performance: VIII Final Report* (Tech. Report HFL-83-1). Blacksburg, VA: Virginia Polytechnic Institute and State University.

A summary report of a five year research program which studied human performance using hard and soft-copy digitally derived imagery. Quality metrics were correlated with human performance results.

Snyder, H. L. (1985). Image quality: Measures and visual performance. In L. E. Tannas (Ed.), *Flat-panel displays and CRTs* (pp. 70- 90). New York: Van Nostrand.

Useful operational definitions of image quality are discussed. Alternative concepts of image quality are offered, mathematical definitions of the various image quality metrics are stated, and results that relate these mathematical quantities to the performance of the user are summarized. The goal is to determine which image quality descriptors or models are valid and meaningful.

Snyder, H. L., Beamon, W. S., Gutmann, J. C., & Dunsker, E. D. (1980). *An evaluation of the effect of spot wobble upon observer performance with raster scan displays* (AMRL-TR-79-81). Wright-Patterson Air Force Base, OH: Air Force Aerospace Medical Research Laboratory.

Summarizes the development of image quality measures for television and photography. Gives judgements as to their validity. Provides experimental results relating measures of image quality to operator performance from line-scan displays. Shows the utility of the MTF and SNRD as image quality metrics.

Snyder, H. L., Keese, R. L., Beamon, W. S., & Aschenbach, J. R. (1974). *Visual search and image quality* (Tech. Report AMRL-TR-73-114). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.

Several experiments were conducted to evaluate alternate unitary measures of video line scan system image quality. An MTF-based metric was shown to predict well the average effects of several imaging system parameters upon the ability of observers to extract information from both dynamic and static images.

Snyder, H. L. & Maddox, M. E. (1978). *Information transfer from computer-generated, dot-matrix displays* (Tech. Report HFL-78-3). Virginia Polytechnic Institute and State University, Blacksburg, Va.

This report investigated the effects of numerous design parameters of alphanumeric dot-matrix displays upon operator performance. Among the parameters investigated experimentally are dot size, dot shape, dot contrast, dot spacing, matrix size, character size, word context, ambient illuminance, dot luminance, and character font. Operator performance in reading and search tasks was predicted by a linear regression model and subsequently cross-validated by additional experiments.

Snyder, H. L., & Maddox, M. E. (1980). On the image quality of dot-matrix displays. In *Proceedings of the Society for Information Display*, 21, (pp. 3-7). New York: Palisades Institute.

This article summarizes the results of a three-year research project which investigated the image quality of dot-matrix displays. Design recommendations are reported and future research needs are noted.

Snyder, H. L., Shedivy, D. I., & Maddox, M. E. (1980). *Quality metrics of digitally derived imagery and their relation to interpreter performance: III. Subjective scaling of hard-copy digital imagery*(Tech. Report HFL-81-3). Blacksburg, VA: Virginia Polytechnic Institute and State University.

This study evaluated subjective image quality of hard-copy digital imagery. Trained photointerpreters judged the interpretability of scenes that were degraded by noise and blur. The NATO scale was revised for use in this research. Results indicated that the different

levels of noise, blur, and the scene content affect the judged interpretability. Analysis also indicated that at least 62 categories should be used to scale interpretability.

- Snyder, H. L. & Taylor, D. F. (1976). *Computerized analysis of eye movements during static display visual search* (Tech. Report AMRL-TR-75-91). Wright-Patterson Air Force Base, Ohio: Aerospace Medical Research Laboratory.

Fixation duration, interfixation distance and number of eye fixations were measured while subjects searched a static display for one target. Non target density was varied. Fixation duration was unaffected by density. Interfixation distance decreased linearly with increases in nontarget density. The authors concluded that the decrease in interfixation distance resulted in longer search times as nontarget density increased, and in increased numbers of fixations per trial.

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

This study manipulated four display parameters to evaluate the sensitivity of four different response measures; accuracy, response time, tachistoscopic recognition, and threshold visibility. Response accuracy was determined to be the most sensitive measure.

- Snyder, H. L., Turpin, J. A., & Maddox, M. E. (1980). *Quality metrics of digitally derived imagery and their relation to interpreter performance: II. Effects of blur and noise on hard-copy interpretability*. (Tech. Report HFL 81-2). Blacksburg, VA: Virginia Polytechnic Institute and State University.

This study evaluated the effects of blur and noise on an information extraction task using hard-copy digital images. Trained photointerpreters were asked to extract information from the images. The effect of noise was found to be significant. The data in this study was found to correlate well ($r = 0.898$) with subjective ratings of the same scenes.

- Spiker, A., Rogers, S. P., & Cicinelli, J. (1984). Color and brightness contrast effects in CRT displays. *S.I.D. International Symposium Digest of Technical Papers, XV*, 62-64.

In this study eight foreground colors and eight background colors were factorially combined to yield 64 stimuli. Subjects were asked to identify the foreground color and accuracy data were collected. The data were analyzed according to confusions among foreground colors. Colors which were frequently confused were changed and the experiment was repeated. The overall error rate decreased by 7% for experiment 2. Color combinations most frequently confused were identified. The authors give luminance values and chromaticity coordinates for the colors used.

- Stein, I. H. (1980). The effect of active area on the legibility of dot-matrix displays. In *Proceedings of the Society for Information Display*, 21, (pp. 17-20). New York: Palisades Institute.

Stein evaluated the effect of percent active area on character legibility and found that under normal or optimal conditions there is little effect of active area; however, under stressed conditions there is a threshold of 30% active area. A percent active area above 30% in stressed conditions does not appear to add to legibility.

- Suen, C. Y., & Shiao, C. (1980). An iterative technique of selecting an optimal 5 X 7 matrix character set for display in computer output systems. In *Proceedings of the Society for Information Display*, 21(pp. 9-16). New York: Palisades Institute.

This article describes a technique for comparing dot-matrix alphanumeric characters to determine the most distinctive set. Eight different measurements are used to eliminate the different character models.

- Sutton, J., & Powers, J. (1984, April). Bringing new technology to an old industry. *Information Display*, 4-8.

A brief discussion of some of the advantages and disadvantages for using flat-panel display technologies in the sign industry.

Suzuki, K., Aoki, F., Ikeda, M., Okada, Y., Zohta, Y., & Ide, K. (1983). High resolution transparent-type a-Si TFT LCDs. *S.I.D. International Symposium Digest of Technical Papers, XIV*, 146-147.

This article discusses the use of an amorphous Si thin film transistors (a-Si TFT) for active matrix addressing of LCDs. Basic display characteristics such as display area, pixel pitch, and others are reported.

Tannas, L.E. (1985). Electroluminescent displays. In L.E. Tannas (Ed.), *Flat-panel displays and CRTs* (pp. 238-288). New York: Van Nostrand Reinhold Co.

This chapter discusses the history of EL displays, theory of operation, and characteristics of the four different types of EL displays. Tannas goes into some depth about the chemistry, physics, and construction of EL displays.

Tannas, L. E., & Goede, W. F. (1978). Flat-panel displays: a critique. *IEEE Spectrum*, 7, 26-32.

This article discusses some of the fundamental problems of flat-panel display technologies including luminous efficiency, matrix addressing, duty cycle and luminance, uniformity and grayscale, color, and cost.

Task, H. L. (1979). *An evaluation and comparison of several measures of image quality for television displays* (Tech. Report AMRL-TR-79-7). Wright-Patterson Air Force Base, OH: Air Force Aerospace Medical Research Laboratory.

This was a major research effort which was designed to determine the correlations between metric values and observer performance in three target detection/recognition studies in which image quality was varied by changing the system MTF. Several different metrics were studied. In general, the MTFA and JND type metrics performed well.

Task, H. L. & Verona, R. W. (1976). *A new measure of television display quality relatable to observer performance* (Tech. Report AMRL-TR-76-73). Wright-Patterson Air Force Base, OH: Air Force Aerospace Medical Research Laboratory.

The GSFP metric is defined as a nonlinear transform of the MTFA to weight the area near the CTF more heavily than the area well above the CTF. Tests of the GSFP produced slightly greater correlations between observer performance measures and GSFP than between MTFA and performance.

Tsuruta, S., Mitsuhashi, K., Ichikawa, S., & Noguchi, K. (1985). Color pixel arrangement evaluation for LC-TV. In *Conference Record of the 1985 International Display Research Conference*, (pp. 24-26). San Diego, CA: Society for Information Display.

In this report the authors performed subjective evaluations of four different possible pixel arrangements for a color LC-TV display. A computer simulation on a raster system was actually used. Authors report that a triangular (RGB) pixel arrangement resulted in the best subjective evaluations.

van Meeteren, A. (1973). *Visual aspects of image intensification*. Soesterberg, The Netherlands: Institute for Perception TNO.

The ICS metric is defined as the system or display MTF cascaded with the visual system MTF or CTF.

Vanderkolk, R. (1976). Dot matrix Symbology. In *Human Factors in dot-matrix displays* (Tech. Report AFFDL-TR-48). Wright Patterson Air Force Base, OH: Air Force Flight Dynamics Laboratory.

Ten display variables were combined in a fractional factorial design to investigate their effects on alphanumeric legibility. The parameters under investigation were; percent active area, symbol definition, contrast, surround luminance, viewing angle, symbol orientation, motion parameters X and Y translation and rotation. Several main effects and interactions were found to be significant. In general, when the legibility was poor the affect of any parameter was amplified.

Vartebedian, A. G. (1970a). Effect of parameters of symbol formation on legibility. *Journal of the Society for Information Display*, 7, 23-26.

This study investigates the effects of symbol generation technique (stroke versus dot), dot shape, and letter orientation on alphanumeric legibility. Vartebedian asserts that 7 x 9 characters with circular dots are superior to other character configurations. Unfortunately there is some confounding in this experiment.

Vartebedian, A. G. (1970b). The design of Visual Displays. *Bell Labs*, 226-231.

This article is a review of other articles published by Vartebedian (1970a and 1971a). Vartebedian asserts that dot-matrix symbols are superior to stroke symbols. This article does not cite any statistical results.

Vartebedian, A. G. (1971a). Legibility of symbols on CRT displays. *Applied Ergonomics*, 2., 130-132.

This article is almost identical to that written in 1970(a).

Vartebedian, A. G. (1971b). The effects of letter size, case, and generation method on CRT display search time. *Human Factors*, 13, 363-368.

The effects of letter size, case, and generation method (stroke versus dot) on subjects ability to perform a word visual search task were investigated. Generation method was confounded with font. The only significant effect of interest was a main effect of case. Lower case words

Verona, R. W., Task, H. L., Arnold, V. C., & Brindle, J. H. (1979). Direct measure of cathode-ray tube (CRT) image quality. In *Proceedings of the Society of Photo-Optical Instrumentation Engineers, Measurements of Optical Radiations, 196*, (pp. 106-113). Bellingham, WA: SPIE.

A direct measuring technique is described for determining the image quality of raster-scanned cathode-ray tube (CRT) displays. The basis for this technique is the modulation transfer function theory and human visual psychophysical data.

Wagner, D. W. (1977). *Color coding an annotated bibliography*(Tech. Report NWC TP 5922). China Lake, CA: Naval Weapons Center.

An annotated bibliography of 57 references related to the use of color for coding information.

Walraven, J. (1985a). Perceptual problems in display imagery. *S.I.D. International Symposium Digest of Technical Papers, XVI*, 192-195.

Very good summary of some of the perceptual problems that may occur with color displays.

Walraven, J. (1985b, January). The colours are not on the display: a survey of non-veridical perceptions that may turn up on a colour display. *Displays*, 35-41.

This article is very similar to Walraven, 1985a. A summary of some of the perceptual problems that may occur with color displays.

Watanabe, H., Hashimoto, S., Yoshida, M., & Shimizu, H. Tsuzuki, A., and Morokawa, A. (1985). Full-color LC video displays using a simple multiplexing method. *S.I.D. International Symposium Digest of Technical Papers, XVI*, 86-87.

This article discusses a prototype full-color LC-TV. Structure and operating characteristics are reported.

Weber, L. F. (1985). Plasma displays. In L. E. Tannas (Ed.), *Flat-panel displays and CRTs* (pp. 332-414). New York: Van Nostrand.

This chapter gives a complete overview of plasma displays. Topics covered include: the history of plasma displays, gas discharge physics, both DC and AC plasma displays (as well as hybrid displays), and the future of plasma displays.

Williams, R. D. (1981). Design considerations for distance viewed dot-matrix displays. In *Proceedings of the Human Factors Society 25th Annual Meeting*(PP.26-29). Santa Monica, CA: Human Factors Society.

This article briefly presents design considerations for dot-matrix CRT displays and extends the findings to distance viewed or large screen displays. Design recommendations are given. The author does not always back-up statements with research examples.

Winkler, R. E. (1979). *Readability of electronic displays*(Tech. Report 79-237T). Manhattan, Kansas: Kansas State University.

This report is a brief literature review of alphanumeric research. Recommendations for several display parameters are given.

Yodoshi, K., Yamaguchi, T., & Niina, T. (1984) A high-brightness monolithic display device using GaP green light-emitting diodes. In *Proceedings of the Society for Information Display, 25*, (pp. 201-206). New York: Palisades Institute.

The fabrication and characteristics of a monolithic LED display are reported.

END

5-87

DTIC