REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

①

AD-A259 665

is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing dat completing and reviewing the collection of information. Send comments regarding this burden estimate or any othe stions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations anc ington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188)

	2. REPORT DATE September 1992	3. REPORT TYPE AND DATE COVERED Book Chapter

4. TITLE AND SUBTITLE Color in Electronic Displays, Chapter 2.1 Research Methods	5. FUNDING NUMBERS None
6. AUTHOR(S) H. L. Snyder and L. J. Trejo	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Navy Personnel Research and Development Center San Diego, California 92152-6800	8. PERFORMING ORGANIZATION REPORT NUMBER BK-92-05
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING Color in Electronics Display, Editors H. Widdell, D. C. Post, and J. Grossman. Plenum Press, New York

11. SUPPLEMENTARY NOTES
Leonard J. Trejo, Navy Personnel Research and Development Center; Harry L. Snyder, Virginia Polytechnic Institute and State University

12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.	12b. DISTRIBUTION CODE A

13. ABSTRACT *(Maximum 200 words)*

This chapter surveys the major research methods and techniques used in the study of color and its effects on human perception and performance. Although a great many research methods have been devised to obtain quantitative data on human vision, only a small subset of those methods are directly pertinent and useful in the study of color sensitivity and the effects of color.

DTIC
ELECTE
JAN 22 1993
S B D

14. SUBJECT TERMS Psychophysical methods, physiological methods, behavioral methods, color discrimination, spectral sensitivity, response time, search time, color coding	15. NUMBER OF PAGES 141
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UNLIMITED

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

COLOR IN ELECTRONIC DISPLAYS

Edited by
HEINO WIDDEL
Forschungsinstitut für Anthropotechnik
Wachtberg-Werthhoven, Germany

and
DAVID L. POST
Armstrong Laboratory
Wright-Patterson Air Force Base, Ohio

Associate Editors:
JEFFREY D. GROSSMAN
Naval Ocean Systems Center
San Diego, California

and
JAN WALRAVEN
Institute for Perception (TNO)
Soesterberg, The Netherlands

PLENUM PRESS • NEW YORK AND LONDON
Published in cooperation with NATO Defense Research Group

44pg

Library of Congress Cataloging-in-Publication Data

Color in electronic displays / edited by Heino Widdel and David L.
Post ; associate editors, Jeffrey D. Grossman and Jan Walraven.
p. cm. -- (Defense research series ; v. 3)
"Published in cooperation with NATO Defense Research Group."
Includes bibliographical references and index.
ISBN 0-306-44191-8
1. Information display systems. 2. Color separation--Data
processing. I. Widdel, Heino. II. Post, David L. III. NATO
Defense Research Group. IV. Series.
TK7882.I6C58 1992
621.39'9--dc20 92-854
CIP

ISBN 0-306-44191-8

A Division of Plenum Publishing Corporation
233 Spring Street, New York, N.Y. 10013

Printed in the United States of America

CHAPTER 2.1

RESEARCH METHODS

Harry L. Snyder

Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061 USA

Leonard J. Trejo

U.S. Navy Personnel Research and Development Center
San Diego, California 92152 USA

PURPOSE

This chapter surveys the major research methods and techniques used in the study of color and its effects on human perception and performance. Although a great many research methods have been devised to obtain quantitative data on human vision, only a small subset of those methods are directly pertinent and useful in the study of color sensitivity and the effects of color.

The more pertinent research methods can be generally classified into psychophysical, physiological, and behavioral methods. Psychophysical methods are those which measure perceptual capabilities of observers and relate the perceptual (psychological) processes to physical dimensions of the stimulus. Psychophysical methods include those that determine the magnitude of sensation attributed to the stimulus. Physiological methods include both central nervous system as well as sensory electrophysiological recordings. Behavioral methods are those which assess the performance capabilities of the observer in performing a task related to the visual stimulus, rather than attempting to measure the perceptual process alone.

PSYCHOPHYSICAL METHODS

This section briefly surveys the more pertinent psychophysical methods used to determine the thresholds of visual perception as well as the magnitude of sensation associated with dimensions of the visual stimulus. In these methods, emphasis is placed on determining the minimum change in the stimulus required to obtain a just-noticeable difference in wavelength, purity, or luminance. Such data are useful in assessing the suitable magnitude of differences used in color coding and color contrast for legibility, as well as to assure that the observer will perceive the displayed color in the fashion intended by the designer.

Color in Electronic Displays, Edited by H. Widdel and D.L. Post
Plenum Press, New York, 1992

Acuity

Acuity is the ability to perceive visual detail clearly. Although acuity is typically measured at either far (6 m) or near (0.4 m) distance, the use of acuity measures in display design is typically limited to the near distance.

To assess acuity, generally high-contrast patterns of various sizes are presented to the observer at a fixed distance, and the observer is asked to determine either the orientation of the pattern (e.g., up, down, left, right) or the pattern itself (e.g., the letters E, S, O). The smallest pattern to be recognized correctly and reliably is defined as the threshold of visual acuity, generally expressed in minutes of visual arc. Visual acuity, as a measure, is then the reciprocal of this angular subtense.

Several forms of visual acuity are described in the experimental literature, including *minimum visible, minimum separable,* and *vernier acuity*. The most common form is that of minimum separable, in which the observer's task is to determine the location of a gap in a circle (i.e., the Landolt ring) or the location of a checkerboard as opposed to a gray square with the same space-average luminance and size. Under optimal conditions, research has shown that the minimum separable acuity ranges from about 30 arcseconds to 1 arcminute. A variety of experimental conditions can influence visual acuity, including target/background contrast, adapting luminance, the nature of the acuity stimulus or target, pupil size, viewing distance, retinal eccentricity of the stimulus, and age of the observer.

Several studies have investigated the influence of stimulus chromaticity on visual acuity, and have found generally that narrowband illumination results in slightly better acuity than does white light illumination of the same intensity. The general conclusion is that the non-color-corrected lens of the eye is better able to focus narrowband illumination than broadband illumination, but this effect has been consistently obtained only for minimum visible and vernier acuity tasks, not for minimum separable tasks (Baker, 1949; Schober & Wittman, 1938; Shlaer, Smith, & Chase, 1942). Interestingly, it has recently been argued that color displays with narrowband monochromatic emissions produce less visual fatigue than do those with broadband achromatic (white) emissions. Murch (1982) has addressed the focusing issue and found that accommodation differences are in the predicted direction.

Color Discrimination

Color discrimination refers to the ability of persons to detect chromatic differences in wavelength or purity. There have been several experimental techniques devised for this purpose, using various psychophysical methods for determination of difference thresholds. In a typical experimental situation, the observer is asked to observe two adjacent or nearly adjacent fields, one containing a standard wavelength light and the other an adjustable (comparison) wavelength. Generally, both fields contain light of high purity, that is, composed of a narrow wavelength band, usually controlled through a monochromator. Further, both fields are equivalent initially in both chromatic content and radiance. The wavelength of the comparison field is changed by the experimenter in small steps, typically one nanometer. The observer is then asked to adjust the radiance of the comparison field until a match can no longer be obtained in the two fields.

The purpose of having the observer adjust the radiance for small wavelength differences is because the visual system is differentially sensitive to different wavelengths, and matches can only be made if the observer is permitted to modify the radiance for different wavelengths. Following this procedure, when a match can no longer be made, it is certain that the mismatch is due to the wavelength difference, not the radiance difference. The difference threshold, following this procedure, is stated as the mean wavelength difference, $\Delta\lambda$, in either direction (increasing or decreasing wavelength) required for the observer to no longer obtain a match.

Variations on this method use fields of various sizes, varying mean radiance (or luminance) levels, and varying levels of stimulus purity. The bipartite field has also been used to study color fusion in the context of wavelength discrimination (Sagawa, 1982), in which one eye receives the same wavelength, λ, on both halves of the bipartite field, while the other eye receives the standard wavelength, λ, on the upper half of its bipartite field and a comparison wavelength, $\lambda + \Delta\lambda$, on the lower half. In such studies, it is generally found that the comparison field presented to the one eye only reduces discrimination but that this reduction is independent of luminance. This result is interpreted as evidence for central color fusion.

Numerous experiments on wavelength discrimination, with many variations in method, have been conducted on color-normal persons. In general, the results show nonuniform sensitivity throughout the spectrum, with greatest sensitivity in the regions of 490-500 and 590-600 nm (Pokorny & Smith, 1986, p. 8-28; Wright & Pitt, 1934). Another frequently referenced result is that of MacAdam (1942), who plotted standard deviations of color matches on the CIE 1931 chromaticity diagram, as illustrated in Figure 1. As can be seen, threshold sensitivity is greatest (i.e., the ellipses are smallest) in the red and blue portions of the diagram and less in the green portion. The fact that ellipses of varying size, rather than circles of equal size, are obtained illustrates the CIE 1931 diagram's lack of perceptual uniformity (which it was never intended to have, although some researchers have mistakenly thought otherwise). Therefore, the apparent discrepancies between MacAdam's (1942) data and the previously mentioned researchers' are illusory; if MacAdam's (1942) data are replotted as a function of wavelength, the results are comparable with the others'.

Heterochromatic Brightness Matching

It is well established that the visual system is not equally sensitive to all wavelengths of light, being less sensitive to reds and blues than to yellows and greens. To quantify this relative sensitivity or photopic luminosity curve, V_λ, several techniques have been developed. A variation on one of those techniques, described here because of its extension for display design purposes, is that of heterochromatic brightness matching. In this procedure, the observer adjusts the luminance (or radiance) of a monochromatic test stimulus of wavelength λ to match the brightness of a fixed reference stimulus, typically white. Both test and reference stimuli are usually presented in a foveally viewed bipartite field of a fixed size, generally 2 degrees in diameter.

Using the convention that the luminance of the test stimulus is B and the luminance of the standard (white) stimulus is L, then the ratio B/L is generally greater than unity and increases with the purity of the test stimulus. Stated another way, the brightness of a chromatic stimulus equal in luminance to a white stimulus is greater, and the extent to which the chromatic stimulus appears brighter is an increasing function of its purity. Differences among observers in B/L ratios are common (Wyszecki & Stiles, 1982), but in general the ratio is larger at both ends of the spectrum than in the middle (Booker, 1981).

B/L ratios become important in the design of visual displays if it is desired that all displayed elements have the same ***brightness***, rather than the same luminance. Similarly, if ***brightness*** coding is intended (rather than luminance coding) of displayed elements, then knowledge of B/L ratios is required. Additionally, some techniques in the scaling of color differences, described below, require equally bright colors or consideration of B/L data.

Minimum Border Distinctness

The classic approach to this method is described by Boynton (1979, pp. 253-255). Consider a bipartite field in which one half consists of light having one color and the other half consists of light having another color. If the radiance of one half is adjusted, there is a point at which the border between the two colors becomes minimally distinct. The "strength"

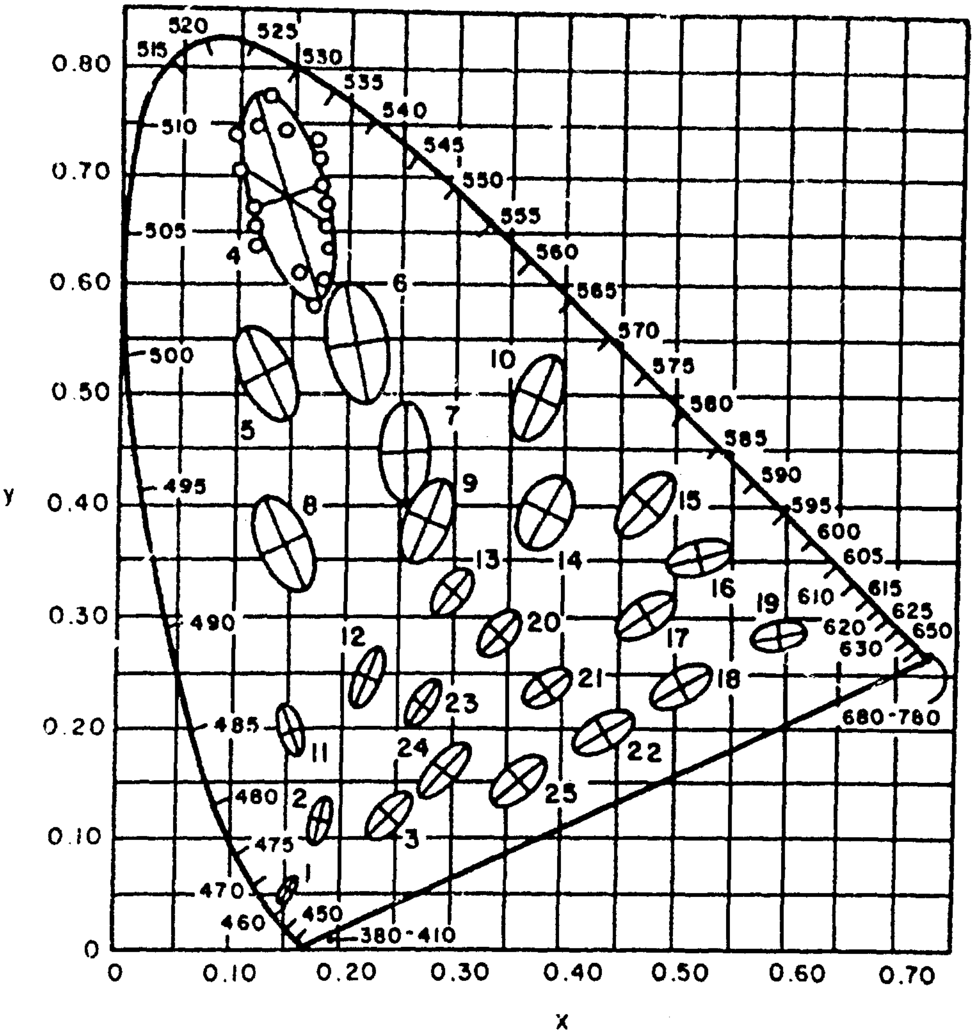

Figure 1. Major and minor axes of ellipses are ten times the standard deviations of color matches. Note that these ellipses might be expected to be noncircular due to (1) the fact that x,y space is not intended to be perceptually uniform, and (2) wavelengths are not equally spaced on the spectrum locus. Redrawn with publisher's permission from MacAdam, D.L. (1942).

of the border's distinctness can then be rated, to obtain a measure of the remaining *chromatic color difference.*

Using this research approach, it has been found that the blue cones contribute less to border discrimination than do the red and green cones (Kaiser & Boynton, 1985), and that the longer wavelengths (red end of the spectrum) are a stronger factor in border discrimination. In fact, stimulus differences in blue-cone stimulation can be very large and yet not produce much or any perceived border. It has also been found that border distinctness is minimized when the luminances of the two colors are equal.

Luminance and Chrominance Modulation Transfer Functions

The application of linear systems analysis and Fourier theory to a variety of systems has similarly influenced the direction of much vision research over the last 25 years. Numerous experiments have been conducted in which a spatial sine-wave grating is used to obtain threshold and suprathreshold visual perception data. The sine-wave grating is of constant luminance and chrominance in one direction, and varies in the orthogonal direction in intensity (or chrominance) in a sinusoidal fashion. The period of the sine-wave is usually related to viewing distance, and is expressed in its reciprocal form as *spatial frequency,* generally in cycles per degree of visual angle so as to make it independent of viewing distance. The amplitude of the sine wave is measured as *modulation,* which is defined as follows:

$$\text{Modulation, } M = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}$$

where L_{max} is the maximum luminance of the sine-wave grating and L_{min} is the minimum luminance of the grating.

In this method, the observer is typically presented with a sine-wave grating at a fixed spatial frequency and with a modulation either well above or well below threshold. Using the psychophysical methods of limits or adjustment, the observer indicates when the modulation of the grating is at threshold for both ascending and descending trials. The mean of a series of such trials at a given spatial frequency defines the threshold for that spatial frequency, while a plot of the means across spatial frequency is termed the modulation threshold function. As described below, the sine-wave grating modulation can be either in luminance, as indicated by the above equation, or in chrominance.

When achromatic sine-wave patterns are used, absolute luminance modulation thresholds are determined as a function of spatial frequency. The U-shaped function is common, with less modulation required at threshold in the region of 2-5 cycles/degree than at either higher or lower spatial frequencies. This function is likened to a bandpass filter, as in electrical systems theory.

The sine-wave grating has also been applied to the determination of chromatic contrast thresholds, in which case the sine-wave grating is held constant in luminance and varied sinusoidally in chrominance about some known chromaticity. The observer is asked to vary the modulation of the sine-wave to determine that modulation which is barely detectable, that is, the threshold. A simple definition of chrominance modulation does not exist, as such a definition depends on the color dimensions in which the stimuli are varied. (This subject is dealt with later in this chapter and throughout this book.)

Threshold chromatic modulation obtained in this fashion is not a bandpass function of spatial frequency, as in the case of luminance thresholds, but rather a low-pass filter function, in which the threshold is fairly constant for frequencies below about 3 cycles/degree and increasing monotonically thereafter. The threshold is higher (more chrominance modulation required) for lower luminance gratings (Van Der Horst, 1969; Van Der Horst & Bouman, 1969). Figure 2 compares modulation threshold functions for both luminance and chrominance modulation.

The chromatic sine-wave grating has also been used to determine wavelength discrimination (Butler & Riggs, 1978; Granger & Heurtley, 1973). In general, the resulting sensitivity function is quite like that found using traditional bipartite fields for wavelength discrimination (e.g., Pokorny & Smith, 1986, p. 8-28). As found by Butler and Riggs (1978), the smallest modulation thresholds are in the regions of 500 and 600 nm, which agree with the data of Wright and Pitt (1934). At these most sensitive wavelengths, the $\Delta\lambda$s are on the order of 1 nm, increasing to approximately 2 nm in the green region (540 nm) and to over 3 nm in the blue (450 nm) and red (650 nm) ends of the spectrum. It should be noted that when chromatic modulation is varied experimentally, as in the Granger and Heurtley (1973) study, by increasing chromatic modulation about a given point on the x,y diagram, there is a concomitant change in excitation purity of the two extremes of the grating. Excitation purity is the proportional distance on the x,y diagram that a color lies between the neutral (white) point and the spectrum locus.

Color Difference Scaling

From some of the above discussion, it is clear that the perceived differences between two visual stimuli are a function of both luminance and chrominance. Since display designers, as well as vision scientists, are often interested in quantifying the extent of the perceived difference between two color stimuli, the issue then arises as to how to scale, into a perceptually uniform volume, the contributions of both color and luminance. To that end, there have been numerous attempts to obtain *uniform chromaticity scales*, in which the distances are linearly related to perceptual differences for colors of the same luminance. In addition, research has been aimed at the definition of *three-dimensional color spaces*, in which both chrominance and luminance differences are considered, such that distances in the three-dimensional space are linearly related to perceived differences between color stimuli differing in both chrominance and luminance. For a thorough discussion of these scales and their derivations see Hunter (1975) and Wyszecki and Stiles (1982).

A widely used attempt at a uniform chromaticity scale for constant-luminance colors was proposed by Judd in 1931, based on the CIE 1931 chromaticity diagram. Subsequent suggestions by MacAdam (1937) and others led the CIE to adopt the 1960 (u,v) diagram and

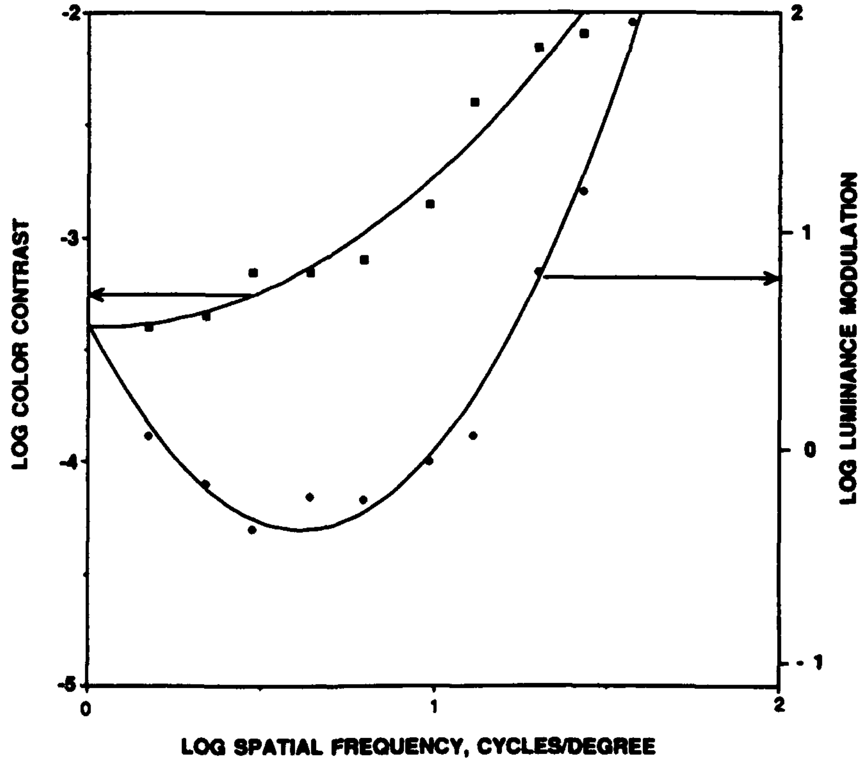

Figure 2. Color contrast is defined as $((\Delta x)^2 + (\Delta y)^2)^{0.5}$, and luminance modulation is as defined in "Luminance and chrominance modulation-transfer functions." Redrawn with publisher's permission from Van Der Horst, G. J. C. (1969).

subsequently, in 1976, the (u',v') diagram, in which $u' = 4X/(X + 15Y + 3Z)$ and $v' = 9Y/(X + 15Y + 3Z)$, where X, Y, and Z are the color's CIE tristimulus values. The CIE 1976 uniform chromaticity scale (UCS) diagram is illustrated in Figure 3.

The typical manner by which a constant-luminance uniform chromaticity scale, such as the 1976 UCS diagram, is converted into a volume to include luminance (or lightness) as well is to add some function of luminance that gives approximately uniform lightness scaling from color stimuli of the same chromaticity. To make the volume perceptually uniform, the scaling of the luminance-related axis must be set in relation to the magnitude of the chromaticity axes. A number of such scales have been advocated, and two have been approved by the CIE, termed CIE 1976 ($L^*u^*v^*$), abbreviated CIELUV, and CIE 1976 ($L^*a^*b^*$), abbreviated CIELAB, in which

$$
\begin{aligned}
L^* &= 116(Y/Y_n)^{1/3} - 16, \\
u^* &= 13L^*(u' - u'_n), \\
v^* &= 13L^*(v' - v'_n), \\
a^* &= 500(X/X_n)^{1/3} - (Y/Y_n)^{1/3}, \text{ and} \\
b^* &= 200(Y/Y_n)^{1/3} - (Z/Z_n)^{1/3},
\end{aligned}
$$

with the constraints X/X_n, Y/Y_n, and $Z/Z_n > 0.008856$, where X, Y, Z are the CIE tristimulus values and u'_n, v'_n, X_n, Y_n, Z_n, are those of a nominally white object-color stimulus. (See Chapter 1.2 for the handling of cases that violate one or more of the

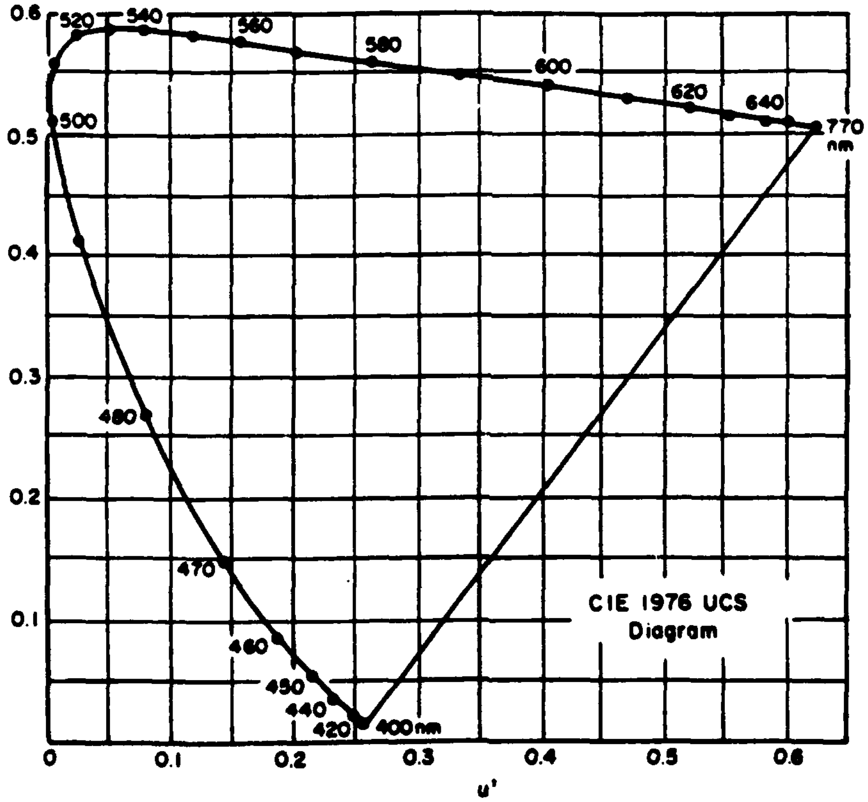

Figure 3. The CIE 1976 UCS diagram.

constraints). The CIE has not defined u'_n, v'_n, X_n, Y_n, Z_n for emissive or self-luminous displays, although various authors have argued for approaches to defining these values so that these formulae might be applied to self-luminous displays (Carter & Carter, 1983; Post, 1984).

Current use of these three-dimensional color spaces includes the assumption that equal distances within the volume are equally perceptible. Then, the perceptual distance (ΔE) between any two colors can be defined by the following formula, for example, in CIELUV:

$$\Delta E = \left[(L_1^* - L_2^*)^2 + (u_1^* - u_2^*)^2 + (v_1^* - v_2^*)^2 \right]^{1/2} ,$$

where L_1^*, L_2^*; u_1^*, u_2^*; v_1^*, v_2^* are the CIELUV coordinates of stimuli 1 and 2, respectively.

The CIELUV and CIELAB spaces are recommended by the CIE for applications involving *object-color stimuli*, such as those encountered in the textile and paint industries. Object-color stimuli consist of "virtually opaque objects illuminated by a light source of given spectral radiant power distribution" (Wyszecki, 1986, p. 9-47). As pointed out by Wyszecki (1986), the CIE has yet to address the uniform scaling of self-luminous color stimuli. Presently, the CIE has a committee working on measurement and scaling of self-luminous displays, but no recommendations have been put forth.

A characteristic of the two CIE 1976 color spaces is that they converge to a narrower range of chromatic differences as L^* decreases. As a result, in some cases, increasing the luminous difference between two stimuli causes their calculated color difference (i.e., ΔE) to decrease and, conversely, decreasing their luminous difference can cause ΔE to increase. This feature is, basically, a consequence of the fact that the CIE 1976 spaces were intended for modeling the perception of reflective surfaces rather than self-luminous ones. For that reason, nonconvergent spaces have been offered and evaluated for use with self-luminous displays. When used on self-luminous displays to predict legibility (see below) of numerals having chromatic and luminance contrast with their background, a nonconvergent space consisting of the dimensions $Yu'v'$ with appropriate scaling outperforms either of the CIE 1976 spaces.

Color-difference scaling in this application to self-luminous displays was obtained from a regression equation using speed of numeral reading as the criterion variable and various luminance and chrominance variables as the predictor variables. For a comparison of alternate predictor spaces see Lippert (1986).

Magnitude Scaling

A variety of scaling methods have been used to obtain estimates of the perceptual strength of the color stimulus. Among the methods that have had the most application to and success in color research are those of *paired comparisons, ratio scaling, magnitude estimation*, and *multidimensional scaling*. In paired comparisons, the stimulus dimension is broken into a number of steps which are estimated to be just below threshold such that the observer has difficulty discerning which of the two members of each pair is greater in magnitude on the stated dimension. Then, the observer is given all possible pairs of the stimuli and asked to select which is greater. Following Thurstone's (1927) *law of comparative judgement, Case V*, the difference between the scale values of any two stimuli is given as the inverse normal transform of the probability that one stimulus is selected over the other. Details of this procedure are given, for example, by Torgerson (1958). The paired comparison technique imposes a relatively simple task upon the observer and thus tends to produce consistent results. It has been used to obtain scales of brightness, saturation, and hue.

Ratio scaling is the technique whereby the observer estimates the magnitude of one stimulus as a multiple or fraction of the magnitude of another stimulus. No adjustment of the stimuli is required by the observer, and the stimuli are generally selected to be noticeably different in magnitude.

In magnitude estimation, the observer is asked to make direct numerical estimates of the perceived magnitude of a set of stimuli, one at a time. In the *modulus* variation of the method, the observer is given the numerical value of one of the stimuli and asked to use it as an anchor from which to determine other ratio scores. In the *free modulus* version, the subject is permitted to select whatever number seems appropriate to him for the first stimulus, and then to use it as the basis for ratio scores for all other stimuli. The instruction to the subject in this version is generally to "call the first stimulus any number that seems appropriate to you. Then assign successive numbers in such a way that they reflect your subjective impression. There is no limit to the range of numbers that you may use" (Stevens, 1975, p. 30).

Based on much research with the method of direct estimation (and other methods), it is generally concluded that the mean estimate, Ψ_x, of the magnitude of a given stimulus attribute increases approximately as the power of the intensity x of the stimulus having that attribute. Thus, for example, the *power law* can be stated as

$$\Psi_x = bx^p.$$

Examples of the application of magnitude estimation and power-law fits can be found for brightness (Marks & Stevens, 1966) and saturation (e.g., Indow & Stevens, 1966). For these and other stimulus dimensions, the power law appears to be a good fit, although variations on its basic form to make the intercept equal to zero (zero physical strength equals zero sensation) are generally recommended (Krantz, 1972).

Multidimensional scaling is an indirect method for estimating the number of the component dimensions, or attributes, that are evoked by a given set of stimuli. Use of this method (e.g., Torgerson, 1958) also is helpful in identifying the nature of the dimensions as well as the relative contributions of the dimensions in eliciting the evoked response strength or perception. The method has had significant benefit in understanding some complex visual problems, such as the components of perceived image quality in photographs. However, in the assessment of color stimuli, it is readily accepted and well proven that the color stimulus consists of three basic perceptual dimensions: hue, saturation, and brightness. Thus, the application of multidimensional scaling is more of academic than applied interest.

PHYSIOLOGICAL METHODS

This section describes physiological research methods useful for studying the effects of electronic display color on the human observer. Most of the research on color that has been performed using physiological methods has addressed basic issues concerning visual function. Another branch of physiological research has dealt with clinical diagnostic procedures. Zrenner (1983) has discussed many basic and clinical applications of physiological methods to primate color vision. The least developed branch of physiological research on color is that which deals with human performance. However, physiological methods offer three distinct advantages over behavioral and psychophysical research methods.

First, physiological measures are directly related to processing of visual information by the nervous system. For this reason they may reflect the operation of mechanisms that intervene between sensory input and behavioral output. Identifying and understanding these mechanisms will allow for better mechanistic models of the neuronal systems that mediate human performance.

Second, unlike behavioral and psychophysical methods, which depend heavily on subjects' knowledge and understanding of experimental manipulations, physiological methods provide simple objective measures of sensory and cognitive function. For example, physiological methods allow for visual sensitivity measurements in infants and children with about the same level of effort as required for adults. This is not generally true for other methods.

Third, physiological methods offer the potential for real-time monitoring of the state of the human operator in complex man-machine systems. This may be of great value in situations where infrequent behavioral responses are required, as in radar monitoring. During behaviorally quiescent periods, physiological measures of brain processing related to probe stimuli on the radar display could provide an estimate of operator attention or alertness.

The application of physiological methods to performance research has not developed sufficiently to allow for a prescription of methods to specific problem areas. Instead the research has dealt with a range of loosely connected problems. Our approach will be to review what we consider to be examples of significant developments in this field. We will supplement this approach by providing enough references to serve as a useful starting point for those interested in using physiological methods in display research.

Much of what we know about the processing of color by the visual system is derived from invasive brain research methods, such as anatomical pathway tracing and microelectrode recordings from neurons in the visual pathways of animals. Because the visual systems of old-world primates such as baboons and macaque monkeys are similar to those of humans, it has been possible to describe the probable basic structure and function of human neuronal mechanisms for processing color. Other, non-invasive brain research methods, such as the electroretinogram (ERG), visually evoked potential (VEP), and the visually evoked magnetic field (VEF), have permitted direct studies of color processing in the human brain. First we will briefly review the structure and function of the primate visual system and then survey the application of non-invasive methods to the study of human color processing.

Brain Mechanisms for Color Processing

Retino-geniculate pathways. The visual image in each eye is first sampled and translated into electrical signals by the photoreceptors, the rods and the three types of cones: long- (R), medium- (G) and short-wavelength sensitive (B). Rods influence color weakly, and these effects are noticeable only in large-field color matching under mesopic conditions or with photopic lights in the orange-red end of the spectrum (Wyszecki & Stiles, 1982, pp. 343-341). Since most electronic displays emit broadband light at photopic levels and present small, foveally viewed symbols, cone signals determine virtually all performance-critical display color processing. Each photoreceptor has a unique spectral sensitivity, the function which relates the absorption rate of photons to their wavelength. These functions are known (Crawford, 1949; Smith & Pokorny, 1975; Vos, 1978; Vos & Walraven, 1971; Wald, 1945) and, for the cones, can be represented by nonlinear combinations of third- or fourth-order polynomial equations (Boynton & Wisowaty, 1980). Figure 4 shows the spectral sensitivity curves of the Smith-Pokorny fundamentals[1].

Individual cone signals are not directly transmitted to the brain. Instead, the basic image-sampling unit in the eye is the ganglion cell[2]. Ganglion cells integrate receptor signals arriving through a network of bipolar, horizontal, and amacrine cells, and transmit them along nerve fibers in the optic nerve to the lateral geniculate nucleus (LGN), a visual input area in the midbrain. In each eye, there are about 1 million ganglion cells, which integrate relayed signals from a variable number of photoreceptors. Most ganglion cells transmit a signal which is directly related to the difference between light falling on a central region (center) and light falling on a surrounding region (surround). Together, these regions are called the receptive field. In a given retinal area, the size and separation of receptive field centers limits the spatial resolving power of the visual system. In the fovea, where acuity is highest, the receptive field centers of ganglion cells apparently receive input from a single cone through midget bipolar cells (Boycott & Dowling, 1969). Proceeding away from the

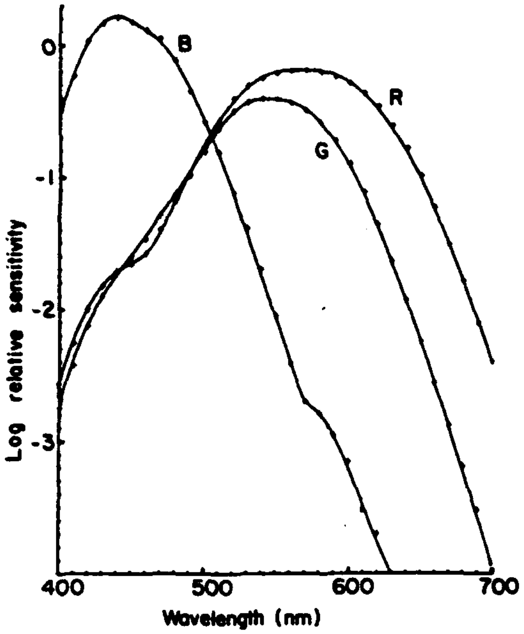

Figure 4. Spectral sensitivity functions of human red-sensitive (R), green-sensitive (G), and blue-sensitive cones (B). Points are derived experimentally (Smith & Pokorny, 1975). Curves are computed from polynomial equations. The relative heights of R and G have been adjusted such that they sum to yield the photopic luminosity function, V_λ. The absolute sensitivities of the three cone types depend on many factors, including individual differences in pigment density, preretinal absorption, retinal eccentricity, and chromatic adaptation. From Boynton and Wisowaty (1980). Copyright by Optical Society of America. Reprinted with permission.

fovea, to retinal areas serving peripheral vision, progressively more photoreceptors—and larger retinal areas—influence the receptive-field centers and surrounds. When the spectral sensitivity of the receptive field center differs from that of the surround, a ganglion cell is said to be spectrally opponent, in addition to the spatial opponency produced by the center-surround organization. Light in the range of wavelengths that stimulates the center produces effects on ganglion cell activity that oppose those produced by light in the wavelength range that stimulates the surround. Spectral opponency is a prerequisite for color coding in the visual system because without it, borders or transients distinguished by wavelength differences could be made invisible to a ganglion cell through intensity adjustments alone. Figure 5 provides a simplified functional diagram of the receptive field structure of a spectrally opponent ganglion cell.

Throughout the retina, most primate ganglion cells are spectrally opponent (Schiller & Malpeli, 1977). Two main types are found: an R-G system, in which signals from the R and G cones oppose each other, and a B-Y system, in which signals from B cones are opposed by either R cones, G cones, or a weighted sum of R and G cones (De Monasterio, Gouras, & Tolhurst, 1975). These opponent-color ganglion cell signals are widely thought to be the physiological basis of human opponent-color mechanisms as expressed by Hering's theory (for a description, see Wyszecki & Stiles, 1982, p. 451).

In addition to the spectrally opponent types, a fraction of primate retinal ganglion cells is considered to be non-opponent or "broadband," and have a spectral sensitivity that closely matches the primate photopic luminosity function (De Monasterio & Schein, 1980). The spectral response properties of these cells suggest that they perform image processing important for achromatic vision.

Since the majority of ganglion cells found in the fovea are both spatially and spectrally opponent, it is thought that they may serve the double duty of both color coding and spatial coding (Ingling & Martinez-Uriegas, 1983). For low spatial frequencies, the ganglion cells transmit color difference signals whereas they transmit luminance-contrast signals for high spatial frequencies. This explains why modulation transfer functions of the human eye differ for luminance and chromatic modulation, the former showing a band-pass characteristic with low frequency attenuation and the latter showing a low-pass characteristic (Kelly & van Norren, 1977), as described above ("Luminance and Chrominance Modulation-Transfer Functions" and Figure 2).

No obvious color transformations are performed by the lateral geniculate nucleus (LGN), the neurons of which directly receive and process spectrally opponent and broadband visual signals from retinal ganglion cells (De Valois, Abramov, & Jacobs, 1966; Kaplan, Purpura, & Shapley, 1987; Schiller & Malpeli, 1977; Wiesel & Hubel, 1966). Instead, it appears that the LGN performs contrast gain control on the retinal ganglion cell signals, which could extend the dynamic range of neurons at higher, cortical levels by preventing response saturation for high-contrast stimuli (Kaplan et al., 1987).

Although obvious color transformations do not occur in the LGN, the structural organization of this nucleus suggests a primary segregation of color-spatial signals from motion- or flicker-sensitive signals. The parvocellular (small-celled) layers contain predominantly color-opponent neurons which linearly combine signals from receptive-field center and surround; the magnocellular (large-celled) layers mostly contain neurons whose spectral sensitivity is "broadband" or at least not clearly color-opponent (Dreher, Fukada, & Rodieck, 1976; Schiller & Malpeli, 1978). About one-third of the magnocellular neurons behave linearly with respect to spatial contrast whereas the remainder exhibit nonlinear spatial interactions (Marrocco, McClurkin, & Young, 1982).

The parvocellular LGN neurons exhibit responses to chromatic modulation which are highly consistent with a dual role in color and spatial vision (Derrington, Krauskopf, & Lennie, 1984). It also appears that transformations of signals from parvocellular neurons by cortical visual neurons can produce the properties of the psychophysical luminance and color opponent channels (Ingling & Martinez-Uriegas, 1983). Signals traversing the magnocellular layers could then subserve other functions, such as flicker and motion detection.

Cortical pathways. Current notions about the structure and function of primate visual cortical areas is covered well in recent reviews (Hubel & Wiesel, 1977; Maunsell & Newsome, 1987; Van Essen & Maunsell, 1983). Here we provide a synopsis of the data relevant to human color vision. The visual cortex is parallel in structure and function, with two major specialized pathways: one for color-spatial vision and one for motion perception. While color is likely to play a role in both of these pathways, the most significant phenomena of human color vision, such as color matching, color or brightness contrast and discrimination, and hue naming, appear to be matched to response properties of neurons in the color-spatial pathway. This system is distinguished from the motion pathway by its anatomical connections, which extend from the primary (striate) visual cortex, V1, through secondary (pre-striate) visual cortex, V2, V3, and V4, to the inferotemporal cortex, IT. Within the color-spatial pathway, there are two major functional systems. One is a system of orientation-selective neurons, which respond best to edges, bars or stripes of a preferred orientation. Another system of neurons is non-oriented; its receptive fields have a concentric organization similar to that found in the retina and the LGN. In V1, neurons of the oriented system are arranged in overlapping patterns of "zebra stripes" in which input from the two eyes alternately dominate cortical activity. Within these stripes the neurons are further

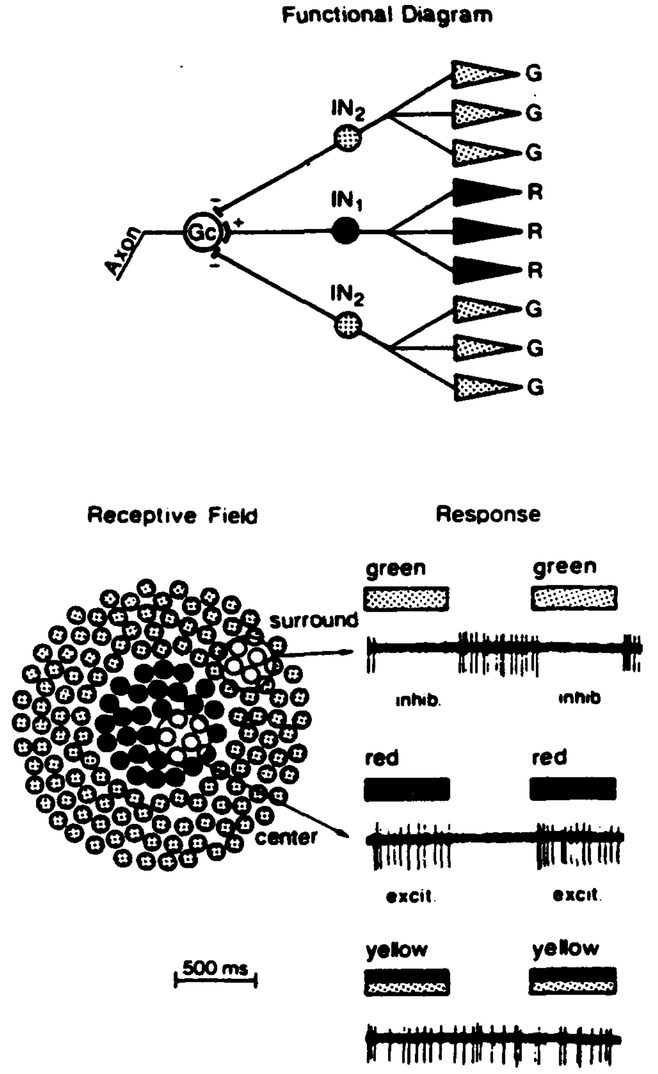

Figure 5. Simplified structure-function diagram for the receptive field of a primate red on-center, green off-surround (+R-G) retinal ganglion cell in the peripheral visual field. Spectral opponency arises from the different spectral sensitivities and influences of center and surround mechanisms. R cone signals from the receptive-field center are integrated by an interneuron, IN_1, which excites the ganglion cell. G cone signals from the surround are integrated by other interneurons, IN_2, which inhibit the ganglion cell. By preferentially stimulating the surround, green light covering the receptive field lowers the ganglion cell firing rate. By preferentially stimulating the center, red light raises the firing rate. A yellow light that equally affects center and surround produces no response. Other types of cells include -R+G, +G-R, -G+R, +B-Y, -B+Y, +Y-B, and -Y+B. From Zrenner (1983). Copyright 1983 by Springer-Verlag. Reprinted with permission.

segregated according to the preferred orientation for an edge, bar, or grating stimulus in the receptive field.

Both oriented and non-oriented V1 neurons exhibit spectral opponency (Michael, 1981). These two neuronal systems appear to divide the task of color-spatial vision into two components. Oriented neurons subserve the identification of objects in terms of edges defined by either color or luminance contrast, whereas non-oriented neurons subserve identification or scaling of hue, saturation, and brightness, and may determine color-contrast effects. Further support for this functional division is given by anatomical studies which have shown that although the oriented and non-oriented systems coexist throughout the color-spatial pathway, they are physically segregated in V1 and V2 (Livingstone & Hubel, 1983, 1984). In V1, neurons in the non-oriented system are clustered together in "blobs" about 0.2 mm wide and spaced about 0.5 mm apart. Neurons from the blob regions provide input to a system of neurons in V2 that is arranged in thin stripes. On either side of each thin stripe is an interstripe region which receives input from the oriented neurons in the inter-blob regions of V1. Beyond the interstripe regions are thick stripes that are part of the motion pathway. V2 neurons in both the thin stripes and interstripe regions in turn project to area V4 (De Yoe & Van Essen, 1985). Here, it has been shown that a significant fraction of neurons are color-coding and exhibit preferences for properties of a stimulus which very nearly follow perceptual phenomena of human color vision. For example, Zeki (1983) stimulated color-coding V4 neurons with light reflected from complex multicolored displays. As the illuminant of this display was varied, these neurons responded selectively to a narrow range of hues, rather independently of the spectrum of the reflected light. This is precisely what is required to allow color constancy of the kind that has been demonstrated in complex scenes when the illuminant is varied (Land, 1974).

Neurons in V4 project to anterior IT which in turn projects to posterior regions of IT. These areas of IT appear to be involved in complex visual functions such as attention, discrimination, and memory. Microelectrode recordings in monkeys during the performance of a delayed matching task have shown that single IT neurons react differentially to stimulus color only when the task requires attention to color (Fuster & Jervey, 1981). Presumably, lower areas, such as V4, perform analyses of stimulus color which allow IT neurons to use color as one of several possible stimulus features that control responding in a complex task.

Physiological Methods for Human Performance Research

Although research with primates has provided clear insight into the cellular mechanisms for human color vision, the invasive techniques—single-cell recordings, anatomical pathway tracing—used in animals have little practical applicability in humans. [However, recordings from retinal ganglion cells in human eyes removed for medical reasons have shown no obvious differences from those of other primates (Weinstein, Hobson, & Baker, 1971).] In humans, a range of non-invasive methods allows a less precise, but in some ways more useful, examination of neural mechanisms involved in color vision as well as of human visual function and more complex visual performance. Here we describe techniques and results of what we consider to be the most important methods: the electroretinogram (ERG), the event-related potential (ERP), the event-related magnetic field (ERF), and pupillometry. Where possible, we cite specific experiments that deal with display or stimulus color as a variable, but in many cases, color display research has not yet capitalized on physiological methods. Therefore, much of the data we describe is not directly applicable to the use of color in electronic displays. However, we think that the relationship of these methods to human visual function and performance is general enough to allow the reader to extrapolate from the results we describe to color display-related problems.

Other potentially useful methods—which have not as yet contributed significantly to color-vision research—are the brain activity mapping techniques: positron emission tomography (PET), nuclear magnetic resonance imaging (NMR), and cerebral blood flow measurements. Although these techniques are providing new insights into brain function, they are limited in temporal resolution, may require injections of drugs or isotopes, and are unlikely to become widely available for research in human performance in the foreseeable future. For these reasons, they will not be considered further here (for reference, see Battistin & Gerstenbrand, 1986).

The electroretinogram. The electroretinogram or ERG is the sum of transient field potentials generated by electrically excitable cells in the retina in response to changes in illumination (Rodieck, 1973). The ERG is typically recorded as the voltage difference between an active electrode (contact lens) placed on the corneal surface and a reference electrode placed away from the eye, typically on the forehead. The interpretation of human ERG waves is complex, but to a first approximation, three major components are observed. In order of increasing latency, they are the a-wave, b-wave, and c-wave. The a- and c-waves appear to directly reflect the activity of retinal neurons, whereas the b-wave probably arises from a depolarization of the (non-neural) Müller cells secondarily to neural activity (Armington, 1974). As indicators of visual function, all of these components have been useful. Steady-state ERG waves are produced when the stimulus is presented repeatedly or flickered with a period shorter than that required for the ERG to resolve to baseline. In this case the components fuse to form a periodic waveform. The amplitude of the steady-state ERG also serves as a gross measure of the effect of a stimulus on retinal activity.

Because the human ERG is strongly influenced by activity throughout the retina, and especially by the rods (Armington, 1974), it is difficult to obtain from it specific cone signals resulting from chromatic stimulation. Since the rod system is slower than the cone system, one way to reduce rod participation in the ERG is to flicker the stimulus at high rates (about 20 Hz or higher). This technique has been used to measure the photopic spectral-sensitivity curve of the human eye (Johnson & Cornsweet, 1954; Padmos & Van Norren, 1971).

Exchange stimulation (also called *silent substitution*) is another method of isolating cone inputs to the ERG and also has wide applicability to the measurement of color visual function (for a detailed description, see Estevez & Spekreijse, 1982). With this method, the spectral radiance of a stimulus is changed as a function of time while other properties, such as size, position, and texture, are held constant. Suitably chosen spectral radiance changes produce differential photon catches in (i.e., isolate) a single cone type while not affecting other cone types. Other spectral radiance changes can isolate any linear combination of receptor types, limited only by the spectral range and number of degrees of freedom of the stimulating device. (Mixing of three independent light sources is required to isolate cone mechanisms.)

Using a variant of the exchange stimulation method, Johnson, Riggs, and Schick (1966) were able to estimate the shape of the wavelength discrimination function of the human eye using the steady-state ERG. Their stimulus was a colored striped pattern (square-wave grating) which was reversed in phase (by displacement) at a rate of 10.7 Hz. The bars of the grating were illuminated by various pairs of monochromatic lights that had previously been matched in brightness. Although their brightness matching method probably did not completely isolate the cones from the rods, a near-isolation was achieved, and fluctuations of stray light in the eye were very small as compared to the wavelength changes. Johnson et al. (1966) found that a wavelength difference of about 20 nm produced a measurable ERG response. In comparison, Riggs and Sternheim (1969) later found that measurable cortical potentials could be produced with much smaller wavelength differences (see below).

More recent research has shown that signals from the cones can also be isolated in the ERG using chromatic adaptation, or by recording the local ERG, an invasive, intraocular variant of the ERG method (Van Norren, 1972; Baron, 1980).

Because of its sensitivity to whole retinal function, the ERG is extensively used in clinical tests for retinal dysfunction. Nevertheless, for research on human performance, it continues to be a difficult and relatively invasive method (corneal electrodes must be sterile and may injure the cornea if improperly applied). For clinical ERG recording, measurement standards have been proposed (Karpe, 1962). Comparable standards for ERG research on human performance have not been proposed.

The event-related potentials. A variety of different and confusing terms and acronyms have been used to refer to the *event-related potentials*. Following Picton (1988), considerable clarity can be gained by defining event-related potentials (ERPs) as the general term for changes in the electric field of the brain (e.g., scalp voltage) that depend on the occurrence of a specific event. When the event has observable physical properties, such as a visual stimulus, and the ERP regularly follows the event in time, it is an *evoked potential*. When the physical properties of the event are undefined or unobservable, such as the absence of an expected stimulus or the occurrence of psychological activity, the ERP is an *emitted potential*.

The event-related potentials are distinguished from the electroencephalogram or EEG by virtue of their relationship to discrete temporal events. In contrast, the EEG is an ongoing, rhythmic variation in the electric field of the brain which is not usually related to specific events. ERPs are further distinguished as being either *exogenous*, i.e., depending primarily on physical events, or *endogenous*, i.e., depending primarily on psychological events.

Both the visually evoked potentials (VEPs), which are exogenous, and other, endogenous ERPs are useful tools for analyzing brain mechanisms of human color vision that subserve human performance with color display systems. Because ERPs are generated by the electric currents surrounding groups of active neurons, they convey information about brain processing of visual stimuli. A distinction between *transient* and *steady-state* evoked potentials is also important in vision and human performance research (Regan, 1988). As for the ERG, steady-state evoked potentials are produced by stimuli that repeat with a period shorter than that required for the evoked potential to resolve to baseline (typically less than 1 s). Transient evoked potentials and ERPs are produced by stimuli or events repeating less frequently or at irregular intervals.

The technology for recording the ERP has advanced significantly in the last 20 years. Modern integrated-circuit amplifiers offer high common-mode rejection, high input-impedance, and low drift, at relatively low cost. Respectively, these amplifier properties provide immunity to environmental electrical noise, enough sensitivity to use contact electrodes which do not pierce the scalp, and stability for recording over extended periods without recalibration. Contact electrodes can now be applied quickly and reliably, using commercially available nylon helmets of varying sizes and electrode configurations. Using inexpensive software and hardware, personal computers now perform ERP signal processing and data recording that previously required expensive clinical equipment or larger computer systems.

In the assessment or prediction of human performance with color display systems, the ERP is a better measure of function than the ERG, for three reasons. First, good-quality ERG recording requires an invasive electrode (direct corneal contact), whereas the ERP does not. Second, for visual stimuli, the ERP arises primarily from the visual cortex, which devotes a disproportionately large fraction of neurons to the central visual field, where color sensitivity is highest. In contrast, the ERG is very sensitive to stray light in the peripheral visual field and to the scotopic visual system. Third, the ERG measures only very peripheral sensory activity whereas the ERP may reflect sensory, perceptual, cognitive, and pre-motor activity.

It is well established that VEP is sensitive to luminance changes. Depending on stimulus size, position, and viewing conditions, luminance transients, step changes, and flicker produce a wide range of variations in waveform amplitude, morphology, and latency (Perry & Childers, 1969; Regan, 1988). Since absolute luminance levels are less important for display quality than luminance contrast, it would be more useful to have an objective measure of response of the visual system to contrast presented by display symbols than of their absolute luminance. Such a measure is provided by the *pattern VEP*, which is typically produced with black and white checkerboards or sine-wave gratings. For patterns defined by luminance contrast, and for which no net change in space-averaged luminance across the pattern occurs, the onset (appearance), offset (disappearance), and contrast reversal of the pattern produce distinct VEP waveforms (see Figure 6). Pattern onset typically produces a sequence of three peaks which alternate in polarity whereas offset typically produces a less complex and less reliable VEP which contains a single peak (Barber, 1984; Spekreijse, Van Der Tweel, & Zuidema, 1973). Pattern reversal differs from pattern onset and offset in that the contrast relationships in the stimulus are reversed, usually suddenly or repetitively, but the pattern itself remains present between reversals. The VEP produced by pattern reversal is generally distinct from that produced by pattern onset but resembles that produced by pattern offset.

The amplitude of the pattern VEP increases linearly with the logarithm of pattern luminance contrast beginning at a point about a factor of two above the psychophysical contrast threshold (Campbell & Maffei, 1970). Depending on the spatial frequency of the stimulus, the linear relation may hold over a range of a factor of three to ten times threshold, beyond which VEP amplitude begins to saturate. Although this relationship has been used primarily to estimate the contrast sensitivity of the visual system by extrapolating to threshold (Regan, 1972; Seiple, Kupersmith, Nelson, & Carr, 1984; Spekreijse, et al., 1973; Tyler & Apkarian, 1985), the same relationship could be used to gauge the relative effect of complex

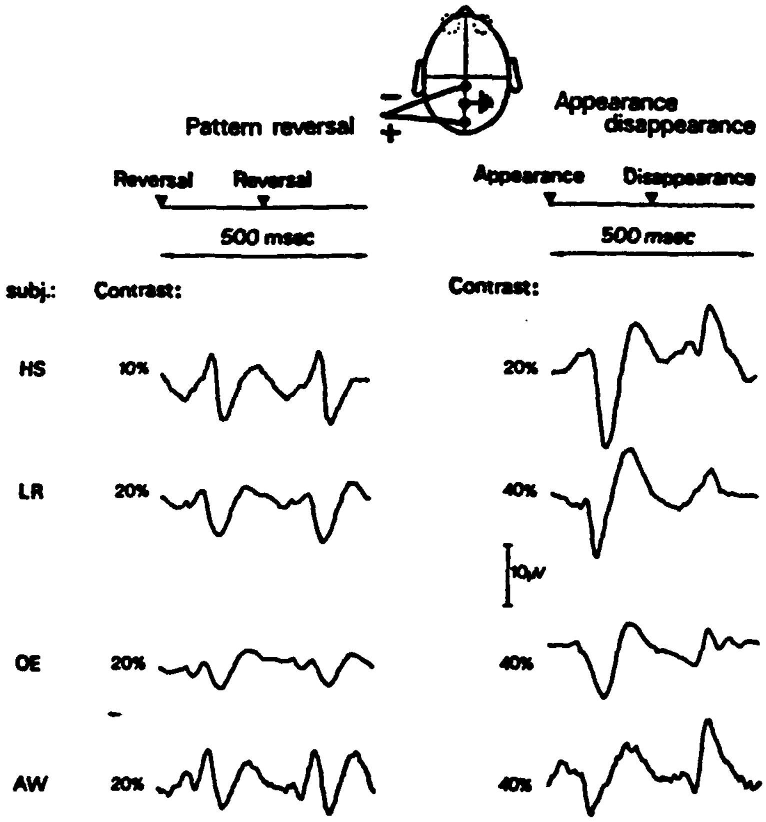

Figure 6. Signal-averaged pattern VEP waveforms of four subjects for reversing (left) and appearing-disappearing (right) checkerboards (3 degrees visual angle, 0.17-degree checks). Space-averaged luminance was held constant; only the contrast of the checkerboard pattern changed at each reversal, appearance, or disappearance. Note that for each subject reversals produce the same waveform pattern whereas appearance and disappearance produce different waveforms. Also note that there are considerable individual differences in the waveforms. From Spekreijse, Van Der Tweel, and Zuidema (1973). Copyright 1973 by Pergamon Press Inc. Reprinted with permission.

display patterns on the human visual system at suprathreshold levels. For example, different complex patterns could be equated for their effect on the visual system by adjusting their contrasts to produce equal pattern-VEP amplitudes.

As for luminance changes, it is well established that color changes and color contrast are effective stimuli for the VEP and pattern VEP. Some of the earliest evidence for this came from an effort to determine whether the VEP could serve as an objective index of luminance matches in heterochromatic flicker photometry (Siegfried, Tepas, Sperling, & Hiss, 1965). Siegfried et al. (1965) found that varying the luminance of a white test field exchanged for a colored standard light in a 3.6-degree central area at a rate of 16 Hz did not always eliminate the steady-state VEP (see Figure 7). In most cases, a residual VEP response persisted at and around the point of minimal subjective flicker. This residual response was later shown to have a rather complex dependence on the temporal frequency of the exchange stimulation and on the particular harmonics of the VEP that are analyzed (Perry, Childers, & Falgout, 1972; Regan, 1970). In particular, Regan (1970), also using a white/colored-light exchange stimulus, showed that the second harmonic of the steady-state VEP for a 24-Hz exchange rate is a sensitive indicator of the contribution of luminance-sensitive mechanisms to the VEP. The minimum in the curve relating the amplitude of the second harmonic to the relative luminance of colored and white lights agreed closely with psychophysical luminance matches. On the other hand, the corresponding curve for the fundamental frequency component showed no clear minimum, thus reflecting sensitivity of the VEP to chromaticity modulation.

Other experiments indicated that subtle variations in the shape of the transient VEP occurred as the wavelength of a brief flash was varied (Shipley, Jones, & Fry, 1965). A portion of this wavelength-related VEP variance appears to arise from differences in the speed with which signals from different cone types reach the visual cortex. Both Krauskopf (1973) and White, Kataoka, and Martin (1977) found that the latency of transient-VEP components in the range between 0 and 250 ms following stimulus onset was greater for stimuli favoring B cones than for stimuli favoring R cones. White et al. (1977) proposed a model in which signals from red- and green-sensitive mechanisms precede those of blue-sensitive mechanisms by about 50 ms in the transient VEP. These findings have potential relevance to the interpretation of human reaction times for colored display symbols or signal lights. However, these studies used stimuli that involved either strong chromatic adaptation or luminance transients and as such, cannot be accepted as proof of variation in conduction latencies for cone signals in the color-spatial pathway.

Still other experiments have shown that it is possible to use the VEP to study the activity of visual pathways when driven by time-varying signals from single cone mechanisms (Estevez, Spekreijse, Van Den Berg, & Cavonius, 1975; Klingaman & Moskowitz-Cook, 1979). Many experiments have shown that higher-order mechanisms are sensitive to pure chromatic contrast (reviewed by Regan, 1988). We cite a few examples below.

As for the ERG, Riggs and Sternheim (1969) found that the steady-state VEP produced by reversal of a pattern of alternating, equal-luminance colored stripes, differing only in wavelength was a sensitive indicator of wavelength differences. They found that VEP amplitude was a nearly linear function of the wavelength difference between the bars in the pattern. By selecting a criterion VEP amplitude for successive wavelength pairs, they correctly estimated the shape of the wavelength discrimination function of the visual system. The slope of the function relating VEP amplitude to wavelength differences was considerably higher than that for the ERG, which indicated higher sensitivity of the VEP to color signals than the ERG.

Using the same principle and a pattern of isoluminant red-green checks which transiently replaced a blank yellow field of the same luminance, Regan (1975) showed that the VEP could be used as an indicator of color deficiency. Such a stimulus produced a clear pattern-onset VEP in a normal observer but not in red-green color-deficient subjects (see Figure 8).

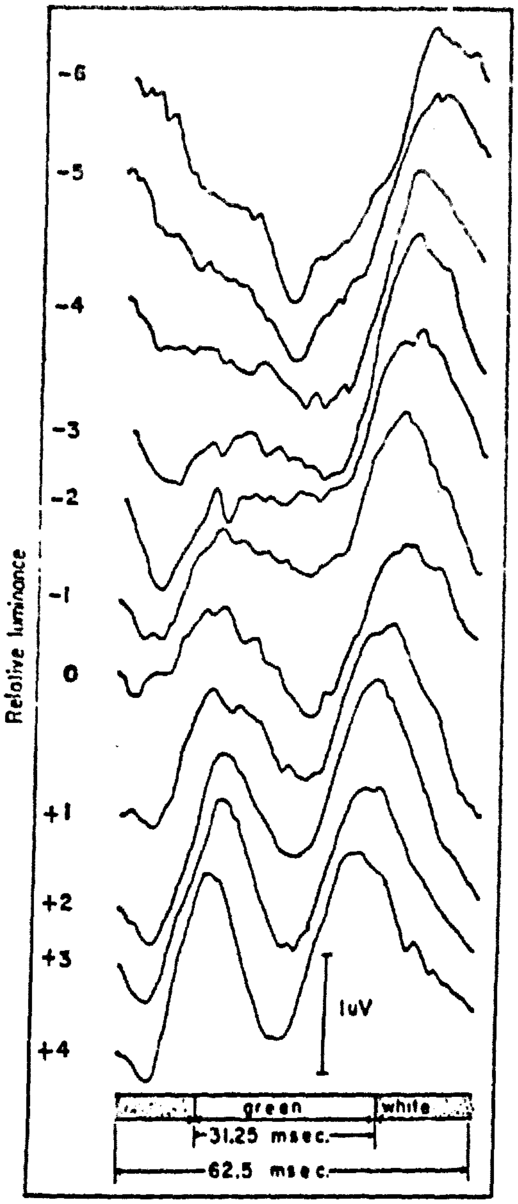

Figure 7. Signal-averaged VEP waveforms for a single subject who viewed a spot that alternated in color (saturation) between white and green. The luminance of the green spot, relative to the point of subjective equality to the white spot, was varied over a range of -6 to +4 steps of 0.07 log-units per step. No luminance setting could be found which eliminated the VEP for the color change. From Siegfried, Tepas, Sperling, and Hiss (1965). Copyright 1965 by American Association for the Advancement of Science. Reprinted with permission.

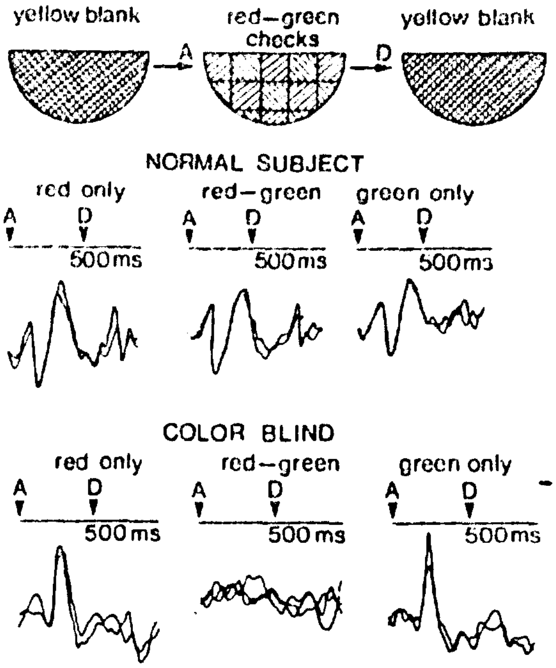

Figure 8. Color pattern VEPs in a normal subject and a red-green color-deficient ("color blind") subject. The stimulus began as a yellow half-circle that abruptly changed to a pattern of checks for 500 ms. As shown in the central column of waveforms, when the checkerboard pattern consisted of isoluminant red and green checks (luminances of the red checks, green checks, and yellow field were all the same), the normal subject gave a clear VEP but the color-deficient subject failed to respond. Luminance-contrast VEPs were normal in both subjects: both the normal subject and the color-deficient subject produced clear VEPs if only the red or green checks were flashed (left and right traces). From Regan (1975). Copyright 1975 by Macmillan Magazines Ltd. Reprinted with permission.

Trejo and Lewis (1988) extended the exchange-stimulation concept to the scaling of color differences at suprathreshold levels for stimuli that either isolated the R-G and B-Y opponent color mechanisms or activated them together in different combinations.Their subjects viewed a 3.5-degree spot that alternated between one color and another of the same luminance at a rate of 1 Hz while the VEP was recorded from pairs of bipolar electrodes over occipital and parietal areas.

Eight pairs of colors were chosen according to two criteria: (1) luminance of all colors was 15 cd/m^2; and (2) total modulation of the opponent-color mechanisms of the human visual system (R-G and B-Y) produced by all pairs was about equal. As shown in Figure 9, not all exchanges produced equal VEP amplitudes. In most subjects, largest VEP amplitudes were found for exchanges of magenta and yellow-green colors. Lowest amplitudes occurred for exchanges of cyan and orange. These VEP results agree with psychophysical data which indicate interactions between the R-G and B-Y mechanisms (Nagy, Eskew, & Boynton, 1987). Combined increases or decreases in B-cone excitation and redness, as in the magenta to yellow-green exchanges, result in summation of signals from the two mechanisms. However, combined increases or decreases in B-cone signals and greenness, as in the cyan-orange exchanges, do not appear to sum. In these cases, the single, most sensitive mechanism appears to determine thresholds. For this reason, chromatic discrimination ellipses, when plotted in a threshold-normalized cone-excitation chromaticity diagram (MacLeod & Boynton, 1979), are elongated at 135 degrees, which corresponds to a cyan/orange axis.

By fitting a symmetrical template to the VEP waveform (see Figure 9), Trejo and Lewis (1988) found that a derived sensitivity measure (reciprocal of ERP amplitude response) predicted the psychophysically observed 135-degree orientation of the chromatic discrimination ellipse (see Figure 10) centered around chromaticity point D_{65}, a CIE standard illuminant which approximates natural daylight (Wyszecki & Stiles, 1982, pp. 144-149). For some subjects, however, the derived sensitivity measure predicted ellipse orientations very different from those found using psychophysical procedures. This suggests that the VEP reflects aspects of color processing in the human visual system other than chromatic discrimination, and emphasizes that VEPs are extremely sensitive to individual differences.

A recent emphasis in ERP research has been on the prediction of human performance of display-related tasks from ERP measures. Basic research has shown that ERPs provide insight into brain mechanisms of cognition (reviewed by Hillyard & Picton, 1987). For visual stimuli, it has been shown that ERP components with latencies less than about 300 ms reflect not only the physical properties of the stimulus, but also reflect the influence of selective attention for the spatial location of the stimulus (Eason, Harter, & White, 1969) and for non-spatial stimulus properties (e.g., color) (reviewed by Harter & Aine, 1984). Other experiments have shown that components with a latency greater than about 300 ms reflect higher-order processes such as the evaluation of stimulus significance, perceptual judgments, and decision making.

Research in engineering and military psychology is beginning to use the predictable relationships between ERP measures and perceptual or cognitive processing to construct models of man-machine interaction. Much early research concerned the inference of mental workload from ERP measures acquired during the performance of complex tasks. These studies (reviewed by Gopher & Donchin, 1986) have shown that the P300 ERP component, a slow positive wave recorded over parietal and central regions of the scalp, may serve as an index of workload when its occurrence depends on probe stimuli that occur during task performance. In general, when the probe stimuli are integral to the primary task (that which has highest priority for the subject), increases in P300 amplitude occur as task difficulty is increased. When the probe stimuli are part of a low-priority secondary task, or are irrelevant, P300 amplitude decreases as task difficulty is increased. These relationships have been observed in a wide variety of tasks, including visual display monitoring (Israel, Wickens, Chesney, & Donchin, 1980) and visuo-motor tracking (Kramer, Wickens, & Donchin, 1983). Further confirmation of these relationships in color simulations of air-defense radar operations has also been reported (Blankenship, Trejo, & Lewis, 1988a, 1988b; Trejo, Lewis, & Blankenship, 1987).

The relationship between P300 amplitude and task workload varies among subjects, and this variation has been linked to individual performance. Using irrelevant visual-probe stimuli and an air-defense radar task, the range over which an RMS-amplitude measure of fronto-central P300 varied as a function of task difficulty was correlated with the average task performance level in a sample of 30 subjects (Trejo, Lewis, & Blankenship, 1989). Subjects who exhibited large P300 RMS-amplitude under low workload and large workload-related decreases in P300-RMS tended to perform better than subjects with lower initial P300-RMS and smaller workload-related decreases. These results are consistent with models of neural information-processing capacity that relate ERP measures to neural resources and their allocation (Defayolle, Dinand, & Gentil, 1971).

The significance of the sensitivity of ERP measures of cognitive processing for the measurement of human performance with display-oriented systems is greatest for situations in which behavioral performance measures are not available. For example, display monitoring tasks may involve long periods during which no measurable responses are

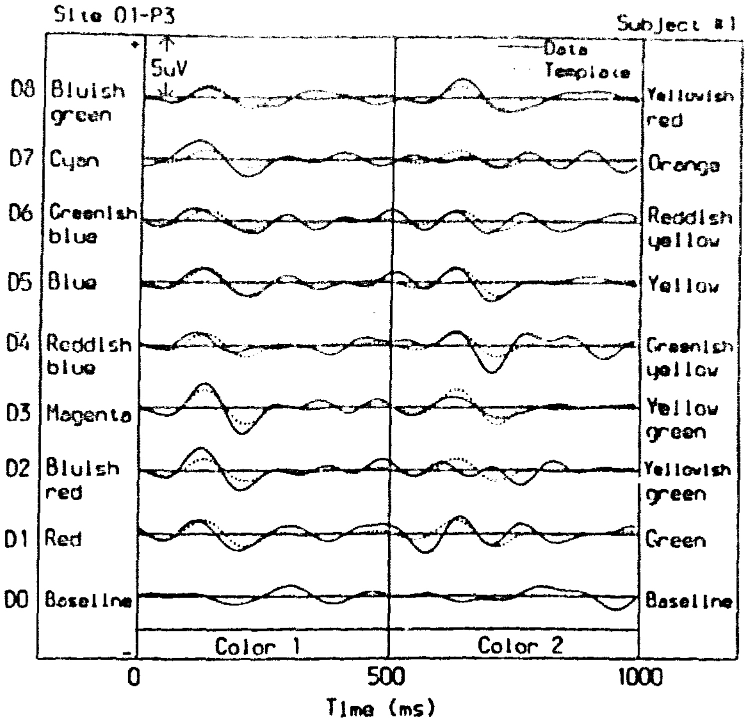

Figure 9. Averaged VEPs derived from the left occipital and parietal areas of one subject for each of eight color-exchange stimuli (D1-D7) and baseline condition (D0) in which no stimulus change occurred. Color-exchange pairs were matched in luminance. For the exchange pairs corresponding to isolation of the R-G (red-green) and B-Y (blue-yellow) mechanisms of opponent-color theory, the color contrast was a constant multiple of the psychophysical threshold. For all pairs, color contrast was such that the absolute value of the sum of R-G and B-Y stimulation was constant. If the mechanisms added together to determine the VEP, the VEPs for all exchanges would be equal in amplitude. Instead, the largest VEPs occurred for exchanges near the magenta/yellow-green pair and smaller VEPs occurred for exchanges near the cyan/orange pair.

required from the operator. During these periods, the intermittent presentation of probe stimuli may allow for ERP-based inferences about the cognitive state of the operator. Such inferences could be useful in comparing the effects of different display color-coding schemes on operator performance.

Although the principles and technology for applying ERPs to analysis of human performance with color displays have existed for some time, little research has been performed to date. In one recent study, ERPs were examined as potential predictors of signal detection and classification performance for stimuli presented on a color CRT (Trejo & Lewis, 1989a, 1989b). Subjects first adapted to a large white background with a luminance of 35 cd/m^2. The background chromaticity was matched to CIE standard illuminant D_{65}. Then, a series of colored stimuli appeared at intervals varying between 1.5 and 2.5 s. The stimuli were squares which replaced the central 7-degrees of the background for 10 ms. Each flash differed from the background in either chromaticity or luminance, but not both. The chromatic stimuli were selected to uniquely activate either the R-G mechanism or the B-Y mechanism, whereas the achromatic stimuli uniquely activated the luminance mechanism. In each series of stimuli, only a single mechanism was activated and individual stimuli had an equal probability of imbalancing the mechanism in opposite directions. For example, in the

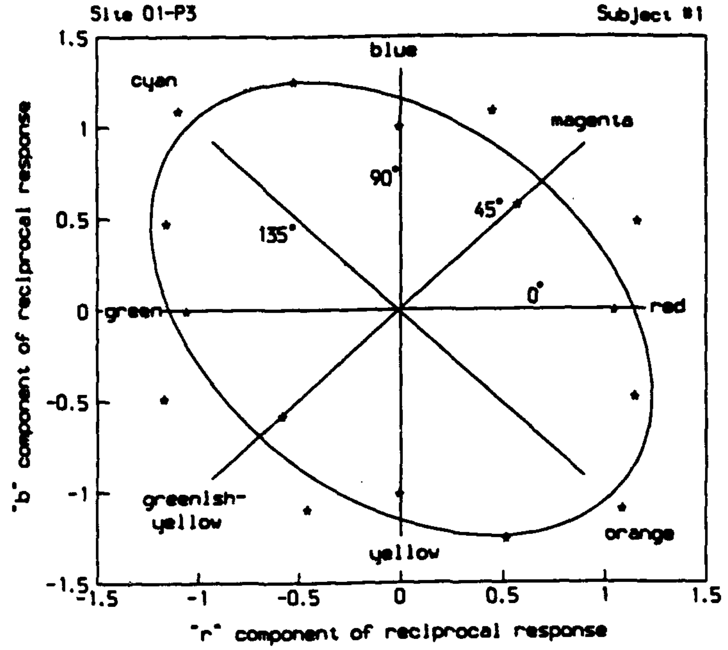

Figure 10. VEP-based color-difference sensitivity derived from the data in Figure 9. A symmetric template was scaled linearly (using least-squares regression) to fit the VEP for each exchange pair. The data points are the reciprocals of the template regression slopes and are plotted as a function of color-exchange direction in a cone-excitation chromaticity plane. At 0 degrees (red-green) and 90 degrees (blue-yellow) the R-G and B-Y mechanisms are isolated. The ellipse of best fit to the reciprocals of the VEP template regression slopes is superimposed on the data points and has its major axis oriented at 134 degrees, near the 135-degree line (D7). The ellipse was fitted using nonlinear least-squares regression with its center constrained to the origin. This orientation agrees with that of chromatic discrimination ellipses reported in comparable psychophysical studies (see text).

R-G series, stimuli appeared either greenish or reddish with respect to the background. In the B-Y series, stimuli appeared either bluish or yellowish and, in the achromatic series, stimuli appeared either black or white.

Subjects were required to detect the stimuli and classify them by pressing one of two keys. Stimuli which appeared as increases in redness (+R-G), blueness (+B-Y), and whiteness (+luminance) were called "targets" and were associated with a "yes" key. Increases in greenness (-R+G), yellowness (-B+Y), and blackness were called "nontargets" and were associated with a "no" key. ERPs for two subjects are shown in Figure 11. The results indicated a direct relationship between detection/classification performance and RMS-amplitude measures of two parietally recorded ERP components. These components were the P300 (see above) and the N1, a sharp negative wave with a latency of approximately 200 ms. When the chromatic contrast of the stimuli was increased, measures of performance accuracy and the two ERP RMS-amplitude measures increased in parallel. Thus, ERP measures may serve as an objective gauge of the effectiveness of color contrasts in providing useful information to the operator of a color display system.

The event-related magnetic field. Recent advances in superconducting materials and technology have permitted the construction of magnetometers sensitive enough to measure

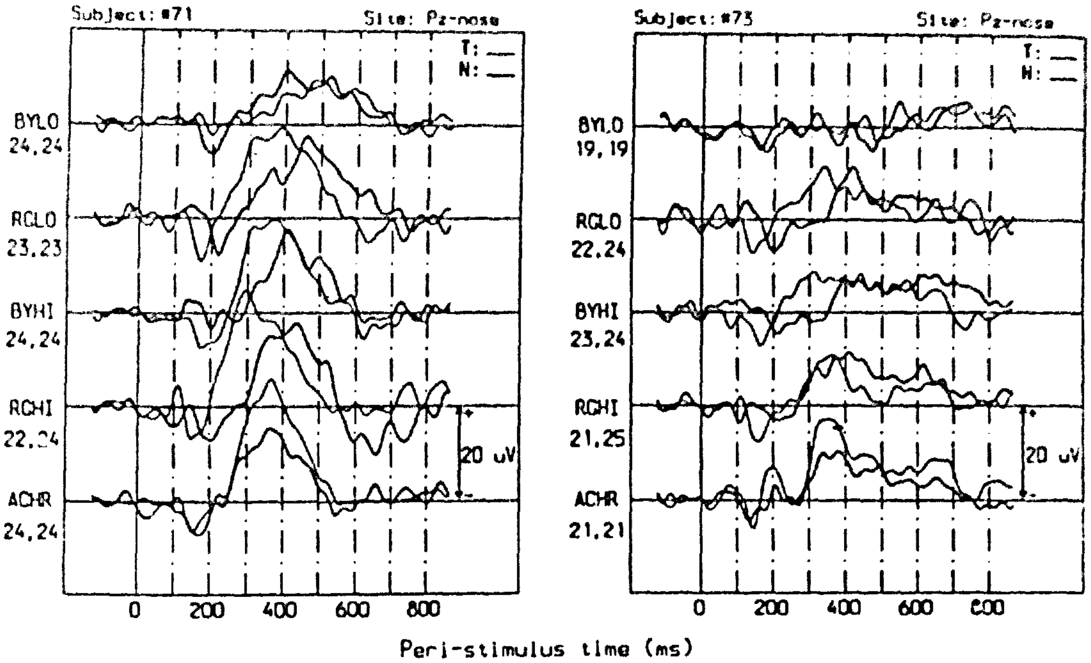

Figure 11. Average ERPs obtained in four different color-contrast conditions and an achromatic-contrast condition in two subjects during the performance of a signal detection and classification task. Numbers of single epochs included in each average ERP appear on the ordinate. Stimulus occurred at time zero on the abscissa. Recording was at the midline parietal site, and the reference electrode was on the nose. Positive voltages are plotted upwards. Heavy traces are for target ("yes") colors: blue, red, and white. Light traces are for nontarget ("no") colors: yellow, green, and black. Details in the text.

magnetic fields on the order of less than 1 picotesla. These advances have permitted direct measurement of brain magnetic fields, which are on the order of 0.1 to 1 picotesla, and thus have given rise to the field of neuromagnetometry (for an introduction to neuromagnetometry see Beatty, Barth, Richer, & Johnson, 1986). The brain generates both a spontaneous time-varying magnetic field, the magnetoencephalogram, (MEG) and event-related magnetic fields (ERFs). It is likely that the most important sources of the MEG and the ERF are groups of parallel, simultaneously active pyramidal cells in the neocortex of the brain (Okada, 1983).

Currently, neuromagnetometers are built around low-temperature superconducting devices called SQUIDs (superconducting quantum interference devices) which require large support systems, liquid-helium cooling, special electronics, and are usually operated in magnetically shielded environments. For these reasons, neuromagnetometers are not widely available for research on human performance. However, new developments in the field of high-temperature superconductors are likely to reduce the size and complexity of support systems for neuromagnetometers, and may result in wider availability.

Neuromagnetometry offers two advantages over electrical recording methods for investigating brain activity. The first advantage concerns accuracy of source localization. Differential electrical resistivities in the tissues between the brain and scalp electrodes (skull, scalp, etc.) and skull openings may distort or displace the electric field, allowing only rough localization of the underlying sources. However, because the skull is virtually transparent to magnetic fields, brain magnetic fields appear to be less distorted and more spatially restricted than brain electric fields and, therefore, simpler to interpret (Kaufman, Okada, Brenner, & Williamson, 1981). For example, Richer, Barth, and Beatty (1983) reported that the visually evoked magnetic field with a latency of 120 ms (M120) was consistent with a simple tangential dipole source lying 12.5 mm below the scalp in the hemisphere opposite to the stimulated visual hemifield. In contrast, the corresponding electric field (N120) was more broadly distributed, was less consistent with a simple dipolar structure, and appeared to be deflected frontally with respect to the dipolar source inferred by the M120 measurements. Thus, magnetometry offers potentially superior localization of the sources underlying specific patterns of brain activity than is possible with electrical recording. However, for complex patterns of activity, which involve multiple brain areas acting simultaneously, this advantage is less clear. This has to do with the inability of EF data to discriminate between single and multiple dipoles or multipole sources (for a discussion of these problems, see Nunez, 1986a, 1986b). The second advantage of neuromagnetometry concerns intrusiveness of the measurements. Unlike electroencephalography, neuromagnetometry does not require physical contact with the subject, since magnetic fields can be sensed at distances on the order of 1 cm from the head.

Due to the limited availability and relative novelty of neuromagnetometers, little research has been performed which directly concerns color vision, displays, and human performance. Corresponding electrical and magnetic measurements of brain processing for visual achromatic patterns have indicated that components of both the VEP and the visually evoked magnetic field (VEF) at latencies of 80 and 120 ms arise from common neural generators on the occipital cortex (Richer et al., 1983). Neuromagnetic measurements of the spatial and temporal modulation-transfer functions of the visual system for achromatic contrast are similar to those obtained with electrical recordings (Okada, Kaufman, Brenner, & Williamson, 1982). Thus, the neuromagnetic method offers a relatively untapped potential for further description of human color vision.

Neuromagnetic research is also beginning to consider the measurement of brain processing related to human performance. In the first study of this kind, amplitude measures of the pattern VEF were related to global on-job performance ratings in two groups of military personnel (Lewis, Trejo, Nunez, Weinberg, & Naitoh, 1988). A group of high-performing subjects exhibited higher average field strength and lower intersubject variation in the VEF than did a group of lower-performing subjects. Later, single-trial analyses showed that trial-to-trial variability in the pattern VEF was also significantly higher in the lower-performing group than in the high-performing group (Lewis, Trejo, Naitoh, Blankenship, & Inlow, 1989). These data indicate that the arithmetic sum and variance of magnetic field strength, within defined intervals after a visual stimulus, may be useful as indices of human performance capabilities. The existence of such physiological indices is predicted by neural models of human information processing that relate neuronal resources and their allocation to the limits of human performance (Defayolle et al., 1971; Trejo, Lewis, & Blankenship, 1990). In principle, there is no technical barrier preventing the application of these methods to the analysis of human performance with color displays.

Pupillometry. The pupillary light reflex (PLR), a constriction of the pupil of the eye in response to increases in retinal illumination, provides a measure of the visual effectiveness of the stimulating light. In a pathway beginning at the retina and extending through the midbrain pretectal and oculomotor nuclei, visual signals influence pupil size through parasympathetic fibers in the oculomotor nerve. It has long been known that luminance or brightness changes influence the PLR (Lowenstein & Loewenfeld, 1969).

Early and more recent experiments have conclusively demonstrated that the PLR is sensitive to wavelength or chromaticity changes (Kohn & Clynes, 1969; Saini & Cohen, 1979). Both rod signals and chromatic aberration have been ruled out as sources for the pupillary response to color (Young & Alpern, 1980). Figure 12 shows that the exchange of one wavelength for another in a 1-degree foveal test field elicits a pupillary constriction which increases in a nearly linear fashion with the difference between the wavelengths. Maximal constrictions of about 0.25 mm occur for wavelength differences of about 160 nm. This is a small constriction, as compared to the luminance-evoked PLR, which can range up to several millimeters in size. These constrictions may be too small for use in human performance research. However, experiments with large stimulus fields may demonstrate larger color PLRs.

Other recent research has shown that visual pathways for pupil control are sensitive to the spatial-frequency spectrum of a visual pattern (Van Der Kraats, Smit, & Slooter, 1977). By exploiting this spatial sensitivity, Slooter and van Norren (1980) predicted human visual acuity from pupillary constrictions produced with appearing-disappearing checkerboard patterns.

The technology for measuring pupil size in real time is relatively simple and inexpensive. Commercial systems are available from several sources, some of which include eye-tracking capability. A basic system, however, requires only a closed-circuit television system and a simple electronic circuit (Green & Maaseidvaag, 1967).

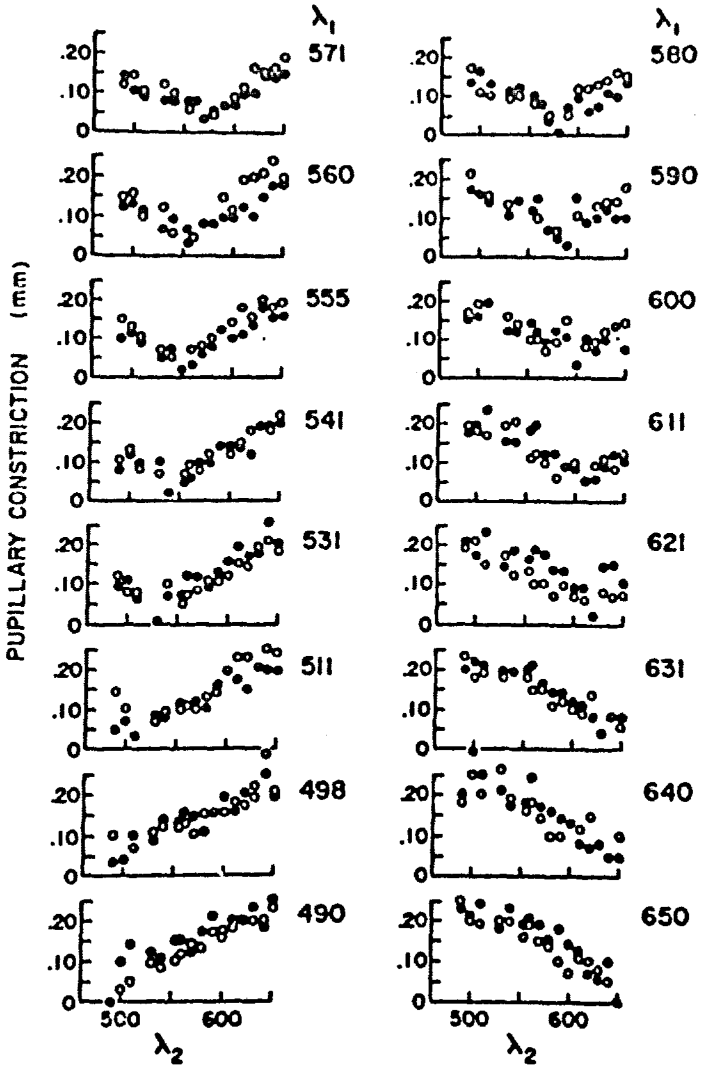

Figure 12. Amplitude of the pupillary constriction produced by exchanges of a variable test wavelength (on the abscissa) for a fixed standard wavelength (on the right) are indicated by the solid circles. Reverse changes (standard to test) are indicated by open circles. Stimulus was a 1-degree spot and luminances of test and standard wavelengths were equal. The size of the PLR increases almost linearly with the wavelength difference between the test and the standard. From Young and Alpern (1980). Copyright 1980 by Optical Society of America. Reprinted with permission.

Because the pupil is also innervated by sympathetic nerve fibers from the superior cervical ganglion, the size of the pupil also reflects the general state of sympathetic arousal. Increased sympathetic arousal produces dilation of the pupil. Through mechanisms that are poorly understood, cognitive processing appears to influence the dilation response of the pupil. A review of these effects (Beatty, 1982) indicates that the pupil dilates in a task-evoked fashion in much the same way that brain potentials are evoked by task-related events. These task-evoked dilations occur with a latency of 100 to 200 ms after the presentation of task stimuli and terminate shortly after task-related processing is complete. Task-evoked pupillary dilations have provided support for neural models of cognitive processing in experiments where psychological workload was manipulated. These tasks have manipulated short-term memory, language processing, reasoning, perception, and sustained or selective attention. In general, the magnitude of the pupillary dilation increases with the level of task difficulty or workload. For example, using visual flashes of an intensity that was seen only 50 percent of the time in a visual signal-detection task, it was shown that the magnitude of pupil dilation was clearly larger for stimuli that were detected than for stimuli that were missed or for blank trials (Hakarem & Sutton, 1966; see Figure 13).

As with neuromagnetometry and the PLR, experiments in which the chromaticity and luminance of visual stimuli are carefully manipulated have not, to our knowledge, been performed with the task-evoked pupillary dilation. However, the method offers the potential for measuring task difficulty or workload manipulations related to display color. Neither do we know of any direct applications of pupillometry to human performance research on color displays, but again there appears to be a large untapped potential. In particular, pupillometry is one of the least invasive physiological techniques, and could conceivably be used for applications such as on-line performance monitoring or for scaling the visual response of individuals to display color or spatial parameters.

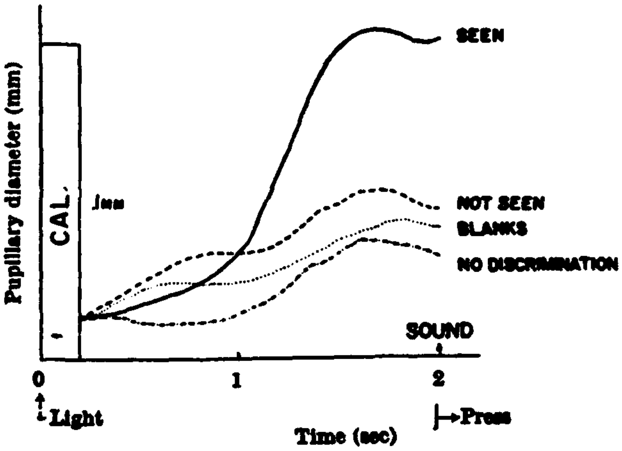

Figure 13. Average task-evoked pupillary dilation for visual stimuli of an intensity that was seen only 50% of the time. There was a clear dilation for lights that were seen as compared to lights that were not seen, blank trials, or trials in which the subject failed to respond (no discrimination). From Hakarem and Sutton (1966). Copyright 1966 by Macmillan Magazines Ltd. Reprinted with permission.

Explaining/Predicting Behavior with Physiological Data

Physiological methods provide at least two ways of explaining human behavior that are distinct from behavioral and psychophysical methods. First, physiological methods measure the sensitivity of the human operator to physical aspects of the display in a way that is relatively free from response biases. For example, the ERG and sensory portions of the ERP and PLR occur within a few hundred ms after a stimulus—usually before any behavioral response—and may be measured in the absence of a behavioral response. Second, physiological methods provide information related to cognitive processes underlying or preceding behavior. Both of the aspects of physiological methods should complement behavioral and psychophysical methods by accounting for process-related variance. Since variability in sensory, perceptual, or cognitive processes leads to variability in psychophysical judgements and behavioral responses, then a physiological measure that is sensitive to those processes will serve to account for variance in performance that is not accounted for by stimulus or task conditions. This principle has been exploited extensively in research on attention, decision making, and cognitive workload (Beatty, 1982; Eason et al., 1969; Gopher & Donchin, 1986), and is now being examined in research with color displays (Trejo & Lewis, 1988). In ERP research, it is now evident that prediction of a behavioral response may be feasible in real-time, even before the response occurs (Gevins, Morgan, Bressler, Cutillo, White, Illes, Greer, & Doyle, 1986).

The utility of physiological data and the information they provide is likely to increase as the power and complexity of electronic displays and systems increases. As the power of display systems increases, the role of the human operator becomes increasingly supervisory. Systems for aircraft navigation, radar, sonar, air-traffic control, and weapons firing can perform many elementary operations automatically. In such contexts, the human operator is susceptible to both cognitive underload and overload. Underload may occur when the system performs so much of the task that the operator loses interest (take, for example, airline pilots who fall asleep when automatic navigation and flight-control systems are in use). Overload may occur when the display presents information at rates greater than the operator can process, leaving the operator with incomplete or inaccurate information about the state of the system. The absence of regular behavioral responses under such conditions makes it difficult to assess the performance level of the operator. Insofar as physiological data correlate with or predict the cognitive states and performance of the operator, they may supplement sparse behavioral data as real-time indicators of operator performance.

BEHAVIORAL METHODS

The methods devised to measure operator performance with color displays range from highly task specific to generic ones that have led to the development of principles for the use (and nonuse) of color. Among the generic methods, the most popular and useful are briefly described here.

Response Time

In many tasks, both in the laboratory and in the field, it is important for the operator to respond quickly and accurately. In some cases, timeliness of response is very important to system performance. For this reason, numerous studies have used response time (e.g., seconds or minutes per response) as the dependent variable to assess the effects of display dimensions and coding.

In a typical experimental scenario, the observer is asked to evaluate some specific display content and then respond by a designated motor or verbal response (e.g., push a button, speak, etc.) The time from the presentation of the stimulus to the time of the response is labeled response time, and it includes all subtasks of visual perception, decision making, and motor response.

Response time to visual stimuli has been used in very simple tasks, such as determining the contribution of chromatic and achromatic activity to simple reaction time (Ueno, Pokorny, & Smith, 1985) as well as very complex visual search and recognition tasks, as will be described below. Usually the response-time data across subjects and/or across experimental conditions are positively skewed, for two reasons. First, there are "floor" limits to the speed with which a subject can respond. Thus, under the best of conditions, all subjects will tend to respond at some lower time limit which represents a minimum perception/decision-making/motor-response cycle. In the case of simple tasks, this total time might be well under one second. Second, more difficult conditions cause lengthening of the response time, which can run into many seconds or minutes. With the increase in difficulty, differences among subjects become more apparent and extraordinarily long times occur. These long times are not necessarily atypical, but merely the result of the sampling from less homogeneous behavior. Thus, a few scores will tend to be very long, causing the positive skewness of the data.

While many experimenters are uneasy about using parametric statistics on skewed data, others recognize the robustness of parametric tests (e.g., F and t) and find no need for a nonlinear transform to reduce or eliminate the skewness. For those researchers who remain uneasy about the skewed distributions of response times, a fortunate solution exists—the reciprocal transform to create the measure of *response speed.* This transform also eliminates the questionable meaningfulness of the mean of a skewed distribution as an estimate of its central tendency.

Response Speed

Response speed is the reciprocal of response time, and is therefore expressed in units of responses per second (or per minute). As such, this measure can eliminate the skewness of response-time data and still remain a meaningful measure per se. (It should be noted that, if the response-time data are normally distributed, then a reciprocal transform of the time data will yield a skewed distribution of speed data.) In many cases, the reciprocal transform is both operationally meaningful and statistically desirable.

Search Time

Visual search time (or speed) has been a favorite dependent variable in much research dealing with color displays. For example, in a review article of color-coding research, Christ (1975) summarized the results of 42 studies published between 1952 and 1973, most of which used search time as the dependent variable. In these studies, the observer is asked to search a visual display for a specific object or class of object, and then to indicate when the object is located and where it is located. The search time begins when the display is presented and ends when the subject responds with some appropriate motor response. The technique has been used successfully to investigate a variety of variables, among them color vs. non-color coding (e.g., Hitt, 1961), number of nontarget objects (Smith, 1962), image quality (Snyder, 1984; Task, 1979), and display failure effects (Abramson, Mason, & Snyder, 1983).

Search time has consistently been shown to be a sensitive variable in such studies and one that relates meaningfully to the observer's task. When both search time and other dependent measures are simultaneously taken, search time tends to correlate highly with the other measures and is often more sensitive to the effects of independent variables. Recent results (Nagy & Sanchez, 1989) indicate that large chrominance or luminance differences between targets and nontargets cause visual search to proceed in a parallel fashion, sampling more than one target at a time, whereas small chrominance or luminance differences produce serial search processes.

Response/Search Accuracy

In some experimental and field situations, users of visual displays will not only show variation in performance (e.g., response or search) times, but also in the number of errors made. Some display designs and environmental conditions are conducive to error generation which, for some systems, is more damaging than are excessive response times. In those cases, a suitable measure of operator performance is response or search accuracy, with the selection of terminology depending on the nature of the task. In general, response accuracy is defined as the number of correct responses divided by the number of total responses, times 100 to obtain a percent measure. Much of the literature on color coding reviewed by Christ (1975) used search accuracy as a response measure, although Christ's (1975) conclusions are essentially the same for search accuracy as they are for search time.

Legibility

Chromatic contrast between a symbol or alphanumeric character and its immediate background can be used to either enhance the ability of a user to locate the symbol (search performance) or to increase the legibility of the symbol by increasing its apparent contrast. Just as increasing achromatic contrast between a symbol and its background will improve the symbol legibility, so will increasing the total effective contrast, the sum of both luminance and chrominance contrast components. Using a method by which chromatic contrast can be transformed into equivalent achromatic (luminance) contrast, it is possible to obtain expressions for total effective contrast between an object and its background (Post, Costanza, & Lippert, 1982).

Studies of legibility for color displays generally present symbols or groups of symbols in a known display location and then ask the subject to read the symbols as quickly and accurately as possible. No search component is present in the task. Rather, the effective contrast between the symbol and its background may be varied, as might other parameters such as symbol size, matrix size, font, strokewidth, etc. The dependent measures for such legibility studies are generally accuracy and response time (or speed).

Response speed has been used in a series of studies to investigate the effect of color contrast on character (numeral) legibility with emissive displays (Lippert, 1986; Post, Lippert, & Snyder, 1983; Sayer, Sebok, & Snyder, 1990). Using numerals or letters of various colors against backgrounds of both uniform color as well as natural scenes, these investigators found that the recommended CIE color spaces (CIELUV and CIELAB) were not as effective in predicting task performance as was a three-space with orthogonal axes of Y, u', and v'. Although additional studies of color spaces are needed to relate different aspects of user performance to color-space dimensions, these studies certainly support the need for additional development of a uniform color space for self-luminous displays.

Color Naming

Color as a coding dimension can have either advantages or disadvantages, depending on its usage (Christ, 1975). When color is used to differentiate classes of objects in a search task, it is critical that the colors used be both perceptually different in an absolute judgment sense, as well as sufficiently unique that they are correctly recognized. Toward this objective, a body of research has dealt with the issue of color naming, that is, the specification of those limits in color space which define the names that users are likely to give to colors. Once such limits or borders for color names are known, then designers can select colors away from the borders so as to minimize color confusions.

Representative of the methodology used in color-naming research is that of Post and Calhoun (1988), in which subjects were asked to use 12 color names on a prepared list and assign one of the names to each color presented on a CRT display. The display color gamut was uniformly sampled from the CIE 1976 UCS (u',v') diagram and displayed in three modes: (1) a solid circle subtending two degrees on a black background; (2) an open square subtending 20 arcminutes on a side on a black background; and (3) an open square on a white (CIE standard illuminant D_{65}) 30-cd/m^2 background. The 12 color names were selected from a prior study which showed that these 12 names represented 88% of all responses which would have been obtained from an unlimited color vocabulary. While the results of these studies are complex and not important in the present context, it is desirable to note that the resulting color-naming boundaries depend on both the reliability with which one wants to avoid overlap of the names (i.e., naming errors) as well as the presentation conditions.

Moreover, recent work by Boynton and his colleagues (e.g., Boynton & Olson, 1987) has demonstrated that the visible colors can be reliably named by the use of 11 basic color names, that these categories are independent of culture and language, and that little or no training is required to obtain consistent color-naming results.

CORRELATIONS AMONG METHODS

It is clear from the preceding material that there are many methods used in the study of the effects of color on observer perception, response, and task performance. The researcher will often select among the alternative measures and methods based on (1) available presentation materials; (2) available recording apparatus; (3) theoretical and scientific theories and interests; and (4) experimental economics. While each of the methods has merit on one or more of these dimensions, it would be desirable to know when it matters, in terms of conclusions to be drawn, which method should be selected and which method may provide less (or more) sensitive results. Unfortunately, few comparative studies exist which illustrate the relative sensitivity of different response measures (but see Snyder & Taylor, 1979, as an exception) to color-display variables. Such studies are needed, both within categories of methods (psychophysical, physiological, and behavioral) and between categories.

Relationships Among Psychophysical Methods

It has long been known that different psychophysical methods (e.g., constant stimuli, limits) can produce different results, particularly for difference thresholds. In fact, precise quantitative comparisons among studies, to be valid, must involve studies that use the same methodology. However, in the area of most absolute-threshold psychophysics, the differences among methods are small but related. That is, different methods may produce slightly larger or smaller threshold values, but the overall relationships are similar.

Relationships Among Behavioral Methods

Such similarities among data using different methods do not frequently hold for behavioral measures, for some behavioral measures are more sensitive than others and yield statistically significant effects of display variables, while other behavioral measures do not produce significant results and have led authors to conclude that such variables have no appreciable effect. In many system design applications, erroneous conclusions of this type can be extremely damaging.

In general, studies have shown that threshold legibility and search performance (time, accuracy) correlate fairly well (Abramson et al., 1983; Kurokawa, Decker, Kelly, & Snyder, 1988; Snyder, 1984; Snyder & Taylor, 1979), although there are also "counterintuitive" examples (Snyder, 1987). At the present time, the state of knowledge in selection of behavioral research methods leans toward using search time and accuracy as the most sensitive and consistent methods to assess the effects of various display variables on user task performance. While differences in sensitivity can be found across experiments using these dependent variables, the differences are minor and rarely misleading. If, however, one chooses to use more subjective, preference-type measures, then significantly different results can be unexpectedly obtained (Snyder, 1987).

Relationships Among Physiological Methods

Correlations among the physiological methods are somewhat stratified. At the lower level are methods that are influenced largely by early sensory processes. Typically, these methods have been compared with psychophysical methods rather than with one another. If two physiological methods yield results that are comparable with a psychophysical method, we may conclude that the two physiological methods are correlated. Using such reasoning, the studies reviewed in this chapter indicate that the PLR, ERG, and exogenous components of the VEP yield similar estimates of spectral sensitivity, wavelength discrimination, and luminance contrast sensitivity. These methods differ mostly in sensitivity, providing either more or less gain. For example, Riggs and Sternheim (1969) found greater voltage changes per unit change in wavelength for the VEP than for the ERG; however, the functions relating voltage change to wavelength change were similar in shape for the two measures.

At the higher level are methods that are sensitive to cognitive, perceptual, and emotional influences. These include the endogenous components of the ERP (e.g., the P300) and the task-evoked pupillary dilation. Few hard data are available concerning the correlations among such methods. This is due in part to the fact that they have been applied to rather complex tasks that have differed greatly across experiments. For example, although the task-evoked pupillary dilation and the P300 component of the ERP both exhibit sensitivity to workload, no experiment that we are aware of has compared these measures to each other or to a third criterion. Presently, it appears that the empirical and theoretical bases for the application of ERP methods to display evaluation are more developed than the corresponding bases for the task-evoked pupillary response. Many more studies have used the ERP than the pupillary response and the diagnosticity of the P300 measure for perceptual/cognitive load versus response selection and execution processes has received empirical support (see Gopher & Donchin, 1986, p. 41-33).

SELECTION OF RESEARCH METHODS

Given the inconsistent and often unpredictable correlations among the various research methods described above, how does one select the method to be used for a particular research question? While a survey of the literature will not answer this question, there are several criteria or guidelines which might be applied. The following might be considered.

Application Relevance

The selected research method should be relevant to the objectives of the experiment. If the objective is to determine basic sensitivity in perceptual response to a display or presentation variable, then one of the psychophysical or physiological methods should be considered. If, for example, one is concerned with perceptual estimates of magnitude of a given stimulus dimension, then direct magnitude estimation (a psychophysical technique) or a visually evoked potential (VEP, a physiological measure) might be used. On the other hand, if the researcher is interested in determining the effect of a given display color-coding scheme on total task (and therefore system) performance, then direct measurement of user response time in a system simulation is more likely to be both valid and meaningful to users. The importance of relevant measures or dependent variables in this sense cannot be overemphasized, for the selection of a sensitive but irrelevant measure can not only be misleading but yield expensively erroneous information.

Reliability

Some measures, particularly electrophysiological ones, are highly repeatable and require only a few samples for accurate estimation of their mean values. Other measures, such as color naming and response accuracy, require large data samples in order to obtain statistical reliability and accurate estimates of population parameters. Only when the researcher is familiar with the reliability of the measure to be used can sample size be selected safely, for selection of too small a sample can lead to conclusions of both insensitivity of the measure as well as poor population-parameter estimates. Of course, requiring too large a sample, while inefficient and costly, cannot lead to inferior population estimates.

Intrusiveness

Many measures, mostly physiological, can be intrusive on the subject and interfere with the behavior or performance of interest. The researcher must be extremely cautious to assure that the measures to be obtained will not interfere with the effect of the independent variables to the extent that the obtained measure is biased or invalid. Again, only a careful analysis of the subject's task in the experiment and the known effects of the experimental method can provide the needed assurances.

Guidelines for Research Control

Throughout the history of research dealing with color displays, there have been examples of otherwise good research that cannot be used because of the lack of experimental control or definition of procedures. Examples of these are numerous, but two of the more critical types are described below to emphasize the point.

Many studies of color coding and its effect have used nominal colors (e.g., "red," "blue," etc.) without precisely specifying their spectral composition, dominant wavelength, or luminance. Because color naming and agreement in color names are not precise, differences in experimental results across these studies may well be explained by differences in control and specificity of the stimuli. To add further uncertainty, numerous studies have inappropriately used the terms "luminance" and "brightness" or have failed to distinguish between these psychophysical and psychological measures. For generalization and understanding of results, it is imperative that displays and visual stimuli used in research be carefully calibrated, preferably radiometrically, so that all stimuli can be stated in spectral power distributions, dominant wavelength, purity, CIE coordinates, and the like. In addition, the nature of the display hardware and software, spatial resolution, addressability, pixel shapes, and the viewing conditions should be stated. Only in this fashion can an experimental study be duplicated in another laboratory.

A second example deals with the popular usage of CRT displays for much of the research of interest. It is not generally appreciated that the red, green, and blue CRT channels for color displays have different (voltage-in/luminance-out) transfer functions which do not maintain constant proportionality across input levels, and that changes in chromaticity may result from changes in luminance unless suitable lookup table conversions are made. Further, there is significant nonlinearity in each of these functions such that simply commanding a greater bit level (or voltage) into the CRT will not result in a proportional increase in luminance (Farley & Gutmann, 1980). For this reason, many of the psychophysical studies of chromatic matching and modulation sensitivity contained in the literature are suspect in their accuracy due to noncalibrated CRTs. Finally, in this regard, it has been shown that even the better CRTs have substantial drift over time, and that a closed-loop beam-current control system is needed to maintain the drift within the difference-threshold limits of the human visual system (Farley, 1987). Without such control, or without very frequent calibration and adjustment of commanded values, threshold data may be suspect.

In summary, regardless of the methodology used, the researcher should be very cautious to calibrate equipment carefully and to maintain calibration through periodic checks. Much of the current sophisticated display equipment is less stable than many researchers realize. In addition, many of the radiometric and photometric devices used for display calibration have substantial noise levels such that their accuracy, particularly per unit bandwidth, is worse than five percent without considerable repetitive sampling and statistical analysis.

FOOTNOTES

1. The Smith and Pokorny fundamentals are linear transformations of Judd's (1951) modification of the CIE 1931 system of primaries, which corrected the photopic luminosity function, V_λ, for excessively low values in the short-wave region of the spectrum (below 460 nm).
2. Many complex lateral interactions occur among retinal cells prior to their integration by ganglion cells, including interactions among neighboring photoreceptors. Nevertheless, the final output of the retina is an ensemble of ganglion cell signals. In this sense, the ganglion cells limit the image information available to the brain, and may be considered as the basic image sampling unit of the visual system.

REFERENCES

Armington, J.C. (1974). *The electroretinogram.* New York: Academic Press.

Abramson, S.R., Mason, L.H., & Snyder, H.L. (1983). The effects of display errors and font styles upon operator performance with a plasma panel. In *Proceedings of the Human Factors Society 27th Annual Meeting* (pp. 28-32). Santa Monica, CA: Human Factors Society.

Baker, K.E. (1949). Some variables influencing vernier acuity. I. Illumination and exposure time. II. Wavelength of illumination. *Journal of the Optical Society of America, 39,* 567-576.

Barber, C. (1984). Convergent features in visual evoked potentials to pattern onset, offset, and reversal. In R.H. Nodar & C. Barber (Eds.), *Evoked potentials II.* Boston: Butterworth.

Baron, W.S. (1980). Cone difference signal in foveal local electroretinogram of primate. *Investigative Ophthalmology and Visual Science, 19,* 1442-1448.

Battistin, L., & Gerstenbrand, F. (1986). *Neurology and neurobiology, 21, PET and NMR. New perspectives in neuroimaging and in clinical neurochemistry.* New York: Alan R. Liss.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91,* 276-292.

Beatty, J., Barth, D.S., Richer, F., & Johnson, R.A. (1986). Neuromagnetometry. In M.G.H. Coles, E. Donchin & S.W. Porges (Eds.), *Psychophysiology.* New York: Guilford Press.

Blankenship, M.H., Trejo, L.J., & Lewis, G.W. (1988a). *Brain activity during tactical decision-making: IV. Event-related potentials as indices of selective attention and cognitive workload* (NPRDC Tech. Note 88-6). San Diego, CA: Navy Personnel Research and Development Center.

Blankenship, M.H., Trejo, L.J., & Lewis, G.W. (1988b). *Brain activity during tactical decision-making: V. A cross-study validation of evoked potentials as indices of cognitive workload* (NPRDC Tech. Note 88-7). San Diego, CA: Navy Personnel Research and Development Center.

Booker, R.L. (1981). Luminance-brightness comparisons of separated circular stimuli. *Journal of the Optical Society of America, 71*, 139-144.

Boycott, B.B., & Dowling, J.E. (1969). Organization of the primate retina: light microscopy. *Philosophical Transactions of the Royal Society of London B, 255*, 109-184.

Boynton, R.M. (1979). *Human color vision*. New York: Holt, Rinehart and Winston.

Boynton, R.M., & Olson, C.X. (1987). Locating basic colors in the OSA space. *Color Research and Application, 12*, 94-105.

Boynton, R.M., & Wisowaty, J.J. (1980). Equations for chromatic discrimination models. *Journal of the Optical Society of America, 70*, 1471-1475.

Butler, T.W., & Riggs, L.A. (1978). Color differences scaled by chromatic modulation sensitivity functions. *Vision Research, 18*, 1407-1416.

Campbell, F.W., & Maffei, L. (1970). Electrophysiological evidence for the existence of orientation and size detectors in the human visual system. *Journal of Physiology, 207*, 637-652.

Carter, R.C., & Carter, E.C. (1983). CIE L*u*v* color difference equations for self-luminous displays. *Color Research and Application, 8*, 252-253.

Christ, R.E. (1975). Review and analysis of color coding research for visual displays. *Human Factors, 17*, 542-570.

Crawford, B.H. (1949). The scotopic visibility function. *Proceedings of the Physical Society of London B, 62*, 321-334.

Defayolle, M., Dinand, J.P., & Gentil, M.T. (1971). Averaged evoked potentials in relation to attitude, mental load and intelligence. In W.T. Singleton, J.G. Fox & D. Whitfield (Eds.), *Measurement of man at work*. New York: Van Nostrand Reinhold Company.

De Monasterio, F.M., Gouras, P., & Tolhurst, D.J. (1975). Trichromatic color opponency in ganglion cells of the rhesus monkey retina. *Journal of Physiology, 251*, 197-217.

De Monasterio, F.M., & Schein, S.J. (1980). Protan-like spectral sensitivity of foveal Y ganglion cells of the retina of macaque monkeys. *Journal of Physiology, 299*, 385-396.

Derrington, A.M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology, 357*, 241-265.

De Valois, R.L., Abramov, I., & Jacobs, G.H. (1966). Analysis of response patterns of LGN cells. *Journal of the Optical Society of America, 56*, 966-977.

De Yoe, E.A., & Van Essen, D.C. (1985). Segregation of efferent connections and receptive field properties in visual area V2 of the macaque. *Nature, 317*, 58-61.

Dreher, B., Fukada, Y., & Rodieck, R.W. (1976). Identification, classification, and anatomical segregation of cells with X-like and Y-like properties in the lateral geniculate nucleus of old-world primates. *Journal of Physiology, 258*, 433-452.

Eason, R.G., Harter, M.R., & White, C.T. (1969). Effects of attention and arousal on visually evoked cortical potentials and reaction time in man. *Physiology and Behavior, 4*, 283-289.

Estevez, O., & Spekreijse, H. (1982). The "silent substitution" method in visual research. *Vision Research, 22*, 681-691.

Estevez, O., Spekreijse, H., Van Den Berg, T.J.T.P., & Cavonius, C.R. (1975). The spectral sensitivities of isolated colour mechanisms determined from contrast EP measurements. *Vision Research, 15*, 1205-1212.

Farley, W.W. (1987). *Design and testing of a luminance and chrominance stabilization system for a computer-controlled color display* (Tech. Report AAMRL-87-027). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

Farley, W.W., & Gutmann, J.C. (1980). *Digital image processing systems and an approach to the display of colors of specific chrominance* (Tech. Report HFL/ONR 80-2). Blacksburg, VA: Virginia Polytechnic Institute and State University.

Fuster, J.M., & Jervey, J.P. (1981). Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science, 212*, 952-954.

Gevins, A.S., Morgan, N.H., Bressler, S.L., Cutillo, B.A., White, R.M., Illes, J., Greer, D.S., & Doyle, J.C. (1986). *Neuroelectric predictors of performance accuracy* (AFOSR Final Report 85-0361). Bolling AFB, D.C.: Air Force Office of Scientific Research.

Gopher, D., & Donchin, E. (1986). Workload—an examination of the concept. In K. Boff, L. Kaufman & J.P. Thomas (Eds.), *Handbook of perception and human performance. Vol. II*. New York: Wiley.

Green, D.G., & Maaseidvaag, F. (1967). Closed circuit television pupillometer. *Journal of the Optical Society of America, 57*, 830-833.

Granger, E.M., & Heurtley, J.C. (1973). Visual chromaticity-modulation transfer function. *Journal of the Optical Society of America, 63*, 1173-1174.

Hakarem, G., & Sutton, S. (1966). Pupillary response at visual threshold. *Nature, 212*, 485-486.

Harter, M.R., & Aine, C.J. (1984). Brain mechanisms of visual selective attention. In R. Parasuraman & D.R. Davies (Eds.), *Varieties of attention*. Orlando, FL: Academic Press.

Hillyard, S.A., & Picton, T.W. (1987). Electrophysiology of cognition. In *Handbook of physiology. Section 1: The nervous system. Volume V: Higher functions of the brain, part 2*. Bethesda, MD: American Physiological Society.

Hitt, W.D. (1961). An evaluation of five different abstract coding methods. *Human Factors, 3*, 120-130.

Hubel, D.H., & Wiesel, T.N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London B, 198*, 1-59.

Hunter, R.S. (1975). *The measurement of appearance*. New York: Wiley.

Indow, T., & Stevens, S.S. (1966). Scaling of saturation and hue. *Perception and Psychophysics, 1*, 253-271.

Ingling, C.R., & Martinez-Uriegas, E. (1983). The relationship between spectral sensitivity and spatial sensitivity for the primate r-g X-channel. *Vision Research, 23*, 1495-1500.

Israel, J.B., Wickens, C.D., Chesney, G.L., & Donchin, E. (1980). The event-related brain potential as an index of display-monitoring workload. *Human Factors, 22*, 211-224.

Johnson, E.P., & Cornsweet, T.N. (1954). Electroretinal photopic sensitivity curves. *Nature, 174*, 614-616.

Johnson, E.P., Riggs, L.A., & Schick, A.M.L. (1966). Photopic retinal potentials evoked by phase alternation of a barred pattern. In H.M. Burian & J.H. Jacobson (Eds.), *Clinical electrophysiology*. Oxford: Pergamon.

Judd, D.B. (1951). Report of U.S. secretariat committee on colorimetry and artificial daylight. In *CIE Proceedings, Volume 1 (7)*, p. 11 (Stockholm, 1951). Paris: Bureau Central de la CIE.

Kaiser, P.K., & Boynton, R.M. (1985). Role of the blue mechanism in wavelength discrimination. *Vision Research, 25*, 523-529.

Kaplan, E., Purpura, K., & Shapley, R.M. (1987). Contrast affects the transmission of visual information through the mammalian lateral geniculate nucleus. *Journal of Physiology, 391*, 267-288.

Karpe, G. (1962). A routine method of clinical electroretinography. *Acta Ophthalmologica Supplement, 70*, 15-31

Kaufman, L., Okada, Y.C., Brenner, D., & Williamson, S.J. (1981). On the relation between somatic evoked potentials and fields. *International Journal of Neuroscience, 15*, 223-239.

Kelly, D.H., & van Norren, D. (1977). Two-band model of heterochromatic flicker. *Journal of the Optical Society of America, 67*, 1081-1091.

Klingaman, R.L., & Moskowitz-Cook, A. (1979). Assessment of the visual acuity of human color mechanisms with the visually evoked potential. *Investigative Ophthalmology and Visual Science, 12*, 1273-1277.

Kohn, M., & Clynes, M. (1969). Color dynamics of the pupil. *Annals of the New York Academy of Science, 156,* 931-950.

Kramer, A.F., Wickens, C.D., & Donchin, E. (1983). An analysis of the processing requirements of a complex perceptual-motor task. *Human Factors, 25,* 597-621.

Krantz, D.H. (1972). Visual scaling. In D. Jameson & L.M. Hurvich (Eds.), *Visual psychophysics*. New York: Springer-Verlag.

Krauskopf, J. (1973). Contributions of the primary chromatic mechanisms to the generation of visual evoked potentials. *Vision Research, 13,* 2289-2298.

Kurokawa, K., Decker, J.J., Kelly, P.L., & Snyder, H.L. (1988). The effects of image rotation on dot-matrix characters. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 1391-1394). Santa Monica, CA: Human Factors Society.

Land, E.H. (1974). The retinex theory of colour vision. *Proceedings of the Royal Institute of Great Britain, 47,* 23-57.

Lewis, G.W., Trejo, L.J., Naitoh, P., Blankenship, M.H., & Inlow, M. (1989, August). Temporal variability of the neuromagnetic evoked field: Implications for human performance assessment (pp. 217-220). In S.S. Williamson, M. Hoke, G. Stroink, & M. Kotani (Eds.), *Advances in biomagnetism*. New York: Plenum.

Lewis, G.W., Trejo, L.J., Nunez, P., Weinberg, H., & Naitoh, P. (1987). Evoked neuromagnetic fields: Possible implications for indexing performance. In K. Atsumi, T. Katila, M. Kotani, S.J. Williamson & S. Ueno (Eds.), *Biomagnetism 1987. Proceedings of the 6th International Conference on Biomagnetism.* Tokyo, Japan.

Lippert, T.M. (1986). Color difference prediction of legibility performance for raster CRT imagery. In *SID Digest* (pp. 86-89). New York, NY: Palisades Institute for Research Services, Inc.

Livingstone, M.S., & Hubel, D.H. (1983). Specificity of cortico-cortico connections in monkey visual system. *Nature, 304,* 531-534.

Livingstone, M.S., & Hubel, D.H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience, 4,* 309-356.

Lowenstein, O., & Loewenfeld, I.E. (1969). The pupil. In H. Davson, (Ed.), *The eye. Vol. 3. Muscular mechanisms*. New York: Academic Press.

MacAdam, D.L. (1937). Projective transformations of ICI color specifications. *Journal of the Optical Society of America, 27,* 294.

MacAdam, D.L. (1942). Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America, 32,* 247-274.

MacLeod, D.I.A., & Boynton, R.M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America, 69,* 1183-1186.

Marks, L.E., & Stevens, J.C. (1966) Individual brightness functions. *Perception and Psychophysics, 1,* 17-24.

Marrocco, R.T., McClurkin, J.W., & Young, R.A. (1982). Spatial summation and conduction latency classification of cells of the lateral geniculate nucleus of macaques. *Journal of Neuroscience, 2,* 1275-1291.

Maunsell, J.H.R., & Newsome, W.T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience, 10,* 363-401.

Michael, C.R. (1981). Columnar organization of color cells in monkey's striate cortex. *Journal of Neurophysiology, 46,* 587-604.

Murch, G.M. (1982). Visual accommodation and convergence to multichromatic information displays. In *SID Digest* (pp. 192-193). New York, NY: Palisades Institute for Research Services, Inc.

Nagy, A.L., Eskew, R.T., & Boynton, R.M. (1987). Analysis of color-matching ellipses in a cone-excitation space. *Journal of the Optical Society of America A, 4,* 756-768.

Nagy, A.L., & Sanchez, R.R. (1989). Interaction of color and luminance differences for optimal visual search. In *SID Digest* (pp. 296-299). New York, NY: Palisades Institute for Research Services, Inc.

Nunez, P.L. (1986a). *Locating sources of the brain's electric and magnetic fields: Some effects of inhomogeneity and multiple sources, with implications for the future* (NPRDC Tech. Note 71-86-12). San Diego, CA: Navy Personnel Research and Development Center.

Nunez, P.L. (1986b). The brain's magnetic field: Some effects of multiple sources on localization methods. *Electroencephalography and Clinical Neurophysiology, 63*, 75-82.

Okada, Y.C. (1983). Neurogenesis of evoked magnetic fields. In S.J. Williamson, G.L. Romani, L. Kaufman & I. Modena (Eds.), *Biomagnetism: An interdisciplinary approach*. New York: Plenum.

Okada, Y.C., Kaufman, L., Brenner, D., & Williamson, S.J. (1982). Modulation transfer function of the human visual system revealed by magnetic field measurements. *Vision Research, 22*, 319-333.

Padmos, P., & Van Norren, D. (1971). Cone spectral sensitivity and chromatic adaptation as revealed by human flicker-electroretinography. *Vision Research, 11*, 27-42.

Perry, N.W., & Childers, D.G. (1969). *The human visual evoked response*. Springfield, IL: Thomas.

Perry, T.W., Childers, D.G., & Falgout, J.C. (1972). Chromatic specificity of the visual evoked response. *Science, 177*, 813-815.

Picton, T.W. (1988). Introduction. In T. W. Picton, (Ed.), *Handbook of electroencephalography and clinical neurophysiology. Revised series. Vol. 3. Human event-related potentials*. Amsterdam: Elsevier.

Pokorny, J., & Smith, V.C. (1986). Colorimetry and color discrimination. In K.R. Boff, L. Kaufman & J.P. Thomas (Eds.), *Handbook of perception and human performance. Volume 1: Sensory processes and perception*. New York: Wiley.

Post, D.L. (1984). CIELUV/CIELAB and self-luminous displays: Another perspective. *Color Research and Application, 9*, 244-245.

Post, D.L., Costanza, E.B., & Lippert, T.M. (1982). Expressions of color contrast as equivalent achromatic contrast. In *Proceedings of the Human Factors Society 26th Annual Meeting* (pp. 581-585). Santa Monica, CA: Human Factors Society.

Post, D.L., & Calhoun, C.S. (1988). Color-name boundaries for equally bright stimuli on a CRT: Phase II. In *SID Digest* (pp. 65-68). New York, NY: Palisades Institute for Research Services, Inc.

Post, D.L., Lippert, T.M., & Snyder, H.L. (1983). Color contrast metrics for head-up displays. In *Proceedings of the Human Factors Society 27th Annual Meeting* (pp. 933-937). Santa Monica, CA: Human Factors Society.

Regan, D. (1970). Objective method of measuring the relative spectral luminosity curve in man. *Journal of the Optical Society of America, 60*, 856-859.

Regan, D. (1972). Evoked potentials to changes in the chromatic contrast and luminance contrast of checkerboard stimulus patterns. In G.B. Arden (Ed.), *The Visual System, 8th International ISCERG Symposium*. New York: Plenum.

Regan, D. (1975). Recent advances in electrical recording from the human brain. *Nature, 253*, 401-407.

Regan, D. (1988). Human visual evoked potentials. In T.W. Picton, (Ed.), *Handbook of electroencephalography and clinical neurophysiology. Revised series. Vol. 3. Human event-related potentials*. Amsterdam: Elsevier.

Richer, F., Barth, D.S., & Beatty, J. (1983). Neuromagnetic localization of the transient visual evoked response to patterned stimulation. *Il Nuovo Cimento, 2*, 420-428.

Riggs, L.A., & Sternheim, C.E. (1969). Human retinal and occipital potentials evoked by changes in the wavelength of the stimulating light. *Journal of the Optical Society of America, 59*, 635-640.

Rodieck, R.W. (1973). *The vertebrate retina.* San Francisco: W.H. Freeman and Company.

Sagawa, K. (1982). Dichoptic color fusion studied with wavelength discrimination. *Vision Research, 22,* 945-952.

Saini, V.D., & Cohen, G.H. (1979). Using color substitution pupil response to expose chromatic mechanisms. *Journal of the Optical Society of America, 69,* 1029-1035.

Sayer, J.R., Sebok, A.L., & Snyder, H.L. (1990). Color difference metrics: Task performance prediction for multichromatic CRT applications as determined by color legibility. In *SID Digest* (pp. 265-268). New York, NY: Palisades Institute for Research Services, Inc.

Schiller, P.H., & Malpeli, J.G. (1977). Properties and tectal projections of monkey retinal ganglion cells. *Journal of Neurophysiology, 40,* 428-445.

Schiller, P.H., & Malpeli, J.G. (1978). Functional specificity of lateral geniculate nucleus laminae of the rhesus monkey. *Journal of Neurophysiology, 41,* 788-797.

Schober, H., & Wittman, K. (1938). Untersuchungen über die Sehschärfe bei verschieden farbigem Licht. *Das Licht: Z praktische Leucht-u.Beleuchtungs-Aufgabe, 8,* 199-201.

Seiple, W.H., Kupersmith, M.J., Nelson, J.I., & Carr, R.E. (1984). The assessment of evoked potential contrast thresholds using real-time retrieval. *Investigative Ophthalmology and Visual Science, 25,* 627-631.

Shipley, T., Jones, R.W., & Fry, A. (1965). Evoked visual potentials and human color vision. *Science, 150,* 1162-1164.

Shlaer, S., Smith, E.L., & Chase, A.M. (1942). Visual acuity and illumination in different spectral regions. *Journal of General Physiology, 25,* 553-569.

Siegfried, J.B., Tepas, D.I., Sperling, H.G., & Hiss, R.H. (1965). Evoked brain potential correlates of psychophysical responses: Heterochromatic flicker photometry. *Science, 149,* 321-323.

Shipley, T., Jones, R.W., & Fry, A. (1965). Evoked visual potentials and human color vision. *Science, 150,* 1162-1164.

Slooter, J.A., & van Norren, D. (1980). Visual acuity measured with pupil responses to checkerboard stimuli. *Investigative Ophthalmology and Visual Science, 19,* 105-108.

Smith, S.L. (1962). Color coding and visual search. *Journal of Experimental Psychology, 64,* 434-440.

Smith, V.C., & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research, 16,* 161-171.

Snyder, H.L. (1984). Image quality: Measures and visual performance. In L.E. Tannas, Jr. (Ed.) *Flat panel and CRT displays.* New York: Van Nostrand.

Snyder, H.L. (1987). Counterintuitive criteria for visual displays. In G. Salvendy, S.L. Sauter & J.J. Hurrell, Jr. (Eds.), *Social, ergonomic and stress aspects of work with computers.* New York: Elsevier.

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.

Spekreijse, H., Van Der Tweel, L.H., & Zuidema, T. (1973). Contrast evoked responses in man. *Vision Research, 13,* 1577-1601.

Stevens, S.S. (1975). *Psychophysics.* New York: Wiley.

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: Aerospace Medical Research Laboratory.

Thurstone, L.L. (1927). A law of comparative judgment. *Psychological Review, 34,* 273-286.

Torgerson, W.S. (1958). *Theory and methods of scaling.* New York: Wiley.

Trejo, L.J., & Lewis, G.W. (1988). Sensitivity to hue differences measured by visual evoked potentials. In D. Eggleson (Ed.), *First Navy Independent Research/Independent Exploratory Development Symposium* (CPIA Publication 492). Laurel, MD: Chemical Propulsion Information Agency.

Trejo, L J., & Lewis, G.W. (1989). Brain mechanisms for human color vision: Implications for display systems. In W.E. Montague (Ed.) *IR/IED FY88 Annual Report* (NPRDC Report AP-89-7). San Diego, CA: Navy Personnel Research and Development Center.

Trejo, L.J., & Lewis, G.W. (1989, October). *Individual differences in classification of transient, isoluminant, chromatic signals: A behavioral and electrophysiological analysis* Paper presented at the Annual Meeting of the Optical Society of America, Orlando, Fl.

Trejo, L.J., Lewis, G.W., & Blankenship, M.H. (1987). *Brain activity during tactical decision-making: II. Probe-evoked potentials and workload* (NPRDC Tech. Note 88-12). San Diego, CA: Navy Personnel Research and Development Center.

Trejo, L.J., Lewis, G.W., & Blankenship, M.H. (1990). *Brain activity during decision-making: III. Relationships between probe-evoked potentials, simulation performance, and task performance* (NPRDC Tech. Note 90-9). San Diego, CA: Navy Personnel Research and Development Center.

Tyler, C.W., & Apkarian, P.A. (1985). Effects of contrast, orientation and binocularity in the pattern evoked potential. *Vision Research, 25,* 755-766.

Ueno, T., Pokorny, J., & Smith, V.C. (1985). Reaction times to chromatic stimuli. *Vision Research, 25,* 1623-1627.

Van Der Horst, G.J.C. (1969). Fourier analysis and color discrimination. *Journal of the Optical Society of America, 59,* 1670-1676.

Van Der Horst, G.J.C., & Bouman, M.A. (1969). Spatio-temporal chromaticity discrimination. *Journal of the Optical Society of America, 59,* 1482-1488.

Van Der Kraats, J., Smit, E.P., & Slooter, J.A. (1977). Objective measurements by the pupil balance method. *Documenta Ophthalmologica Proceedings Series, 14,* 213.

Van Norren, D. (1972). Cone spectral sensitivity studied with an ERG method. In G.B. Arden (Ed.), *The visual system. Neurophysiology, biophysics, and their clinical applications.* New York: Plenum.

Van Essen, D.C., & Maunsell, J.H.R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neurosciences, 6,* 370-375.

Vos, J.J. (1978). Colorimetric and photometric properties of a 2^{o} fundamental observer. *Color Research and Application, 3,* 125-128.

Vos, J.J., & Walraven, P.L. (1971). On the derivation of the foveal receptor primaries. *Vision Research, 11,* 799-818.

Wald, G. (1945). The spectral sensitivity of the human eye; a spectral adaptometer. *Journal of the Optical Society of America, 35,* 187.

Weinstein, G.W., Hobson, R.R., & Baker, F.H. (1971). Extracellular recordings from human retinal ganglion cells. *Science. 171,* 1021-1022.

White, C.T., Kataoka, R., & Martin, J.I. (1977). Colour evoked potentials: Development of a methodology for the analysis of the processes involved in colour vision. In J.E. Desmedt (Ed.), *Visual evoked potentials in man: New developments.* Oxford: Clarendon Press.

Wiesel, T.N., & Hubel, D.H. (1966). Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *Journal of Neurophysiology, 29,* 1115-1156.

Wright, K.D., & Pitt, F.H.G. (1934). Hue-discrimination in normal colour-vision. *Proceedings of the Physical Society of London, 46,* 459-473.

Wyszecki, G. (1986). Color appearance. In K.R. Boff, L. Kaufman & J.P. Thomas (Eds.), *Handbook of perception and human performance. Volume I: Sensory processes and perception.* New York: Wiley.

Wyszecki, G., & Stiles, W.S. (1982). *Color science.* New York: Wiley.

Young. R.S.L., & Alpern, M. (1980). Pupil responses to foveal exchange of monochromatic lights. *Journal of the Optical Society of America, 70,* 697-706.

Zeki, S. (1983). Colour coding in the cerebral cortex: The reaction of cells in monkey visual cortex to wavelengths and colours. *Neuroscience, 9,* 742-765.

er, E. (1983). The spectral properties of the human visual system as revealed by sually evoked cortical potentials (VECP) and psychophysical investigations. In E. ner (Ed.), *Neurophysiological aspects of color vision in primates: Comparative ies on simian retinal ganglion cells and the human visual system, studies of brain nction, volume 9,* New York: Springer-Verlag.