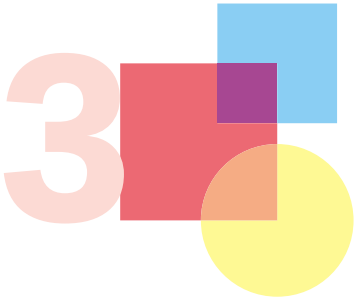


Color Matching and Color Discrimination

Vivianne C. Smith and Joel Pokorny

Departments of Ophthalmology & Visual Science and Psychology, University of Chicago, 940 East Fifty-Seventh Street, Chicago, IL 60637, USA

CHAPTER CONTENTS

The Science of Color
ISBN 0–444–512–519

3.1 INTRODUCTION

The purpose of this chapter is to summarize the data of color matching, detection of spectral lights, and discrimination of lights on the basis of differences in chromaticity. Matching, detection, and discrimination are all tasks in which the spectral content of the lights is an important parameter. The tasks involve either identity matching or detection of a just noticeable difference. Data are expressed in physical units. The color appearance at the identity match or at the just noticeable difference is irrelevant and rarely noted. Although color names are often used to refer to lights as a shortcut form of expression (e.g. red light for a 660 nm spectral light or white light for a continuous spectral power distribution), we have tried to avoid such terminology in this chapter.

The data of matching, detection, and discrimination have been modeled successfully by considering early retinal processing. All the phenomena described in this chapter can be modeled by the activity of parvocellular (PC-), koniocellular (KC-) and magnocellular (MC-) pathways (see Chapter 6). Some readers may find it helpful to read Chapter 6 concurrently with this chapter.

We will concentrate primarily on the classical studies in the literature. These studies have involved small (1° or 2°) or moderate (10°) sized foveally fixated, circular stimuli. Luminance levels of matching and discrimination stimuli are usually at low to medium photopic levels, with retinal illuminations of 5–1000 trolands (td). In detection studies, large concentric backgrounds are often employed and light levels may extend to levels where substantial photopigment bleaching occurs.

3.2 COLOR MIXTURE

A human observer looking at a colored object has no way of knowing from its appearance the spectral composition of the physical stimulus. Colorimetry provides a system of color measurement and specification based upon the concept of equivalent-appearing stimuli. In a color match, two fields of differing spectral radiation appear identical. The match represents a unique neural state in which neural signals generated by the fields are identical.

3.2.1 PRINCIPLES AND PROCEDURES

Metamers: Metameric lights are lights that though of dissimilar spectral radiation are seen as the same by the observer. In a prototypical color-matching experiment using additive lights, the metamers are presented in a bipartite field. For 2° foveal fields, metamers have three important properties that allow treatment of color mixture as a linear system (Grassmann, 1853):

1. The additive property. When a radiation is identically added to both sides of a color mixture field, the metamerism is unchanged.
2. The scalar property. When both sides of the color mixture field are changed in radiance by the same proportion, the metamerism is unchanged.
3. The associative property. A metameric mixture may be substituted for a light without changing the metameric property of the color fields.

According to Grassmann's laws, a color match is invariant under a variety of experimental conditions that may alter the appearance of the matching fields. Metameric matches will hold with the addition of a chromatic surround or following pre-exposure to a moderately bright chromatic field.

Trichromacy: A fundamental property of normal human color vision is the existence of color matches of lights that differ in spectral composition. It is possible to find a metamer for any light (spectral power distribution) by variation of the energies of three fixed lights, which are called primaries. The terms trichromat and trichromacy refer to this property of human color vision. There is wide freedom in the choice of primaries. A formal requirement is that one primary cannot be metameric to a mixture of the other two. In practice, it is desirable that the primaries be spectrally separated as much as possible and that the matching field is at a mid-photopic level. The choice of primaries is dictated largely by experimental convenience. Primaries may or may not themselves be spectral. However, in the development of a colorimetric system, the test lights

should be spectral or near-spectral in order to derive the largest color gamut.

It has become customary to present the results of color mixture experiments as color equations. Suppose we have three primaries, $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$ and a test light, S, arranged in a bipartite field. We find that a mixture of S and $\mathbf{P}_3$ appears identical to a mixture of $\mathbf{P}_1$ and $\mathbf{P}_2$, when the radiant energies of S, $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$ are P_S, $P_{S,1}$, $P_{S,2}$ and $P_{S,3}$ respectively. This is written:

$$P_S S \oplus P_{S,3}\mathbf{P}_3 \equiv P_{S,1}\mathbf{P}_1 \oplus P_{S,2}\mathbf{P}_2 \qquad (3.1)$$

where '≡' means visually identical and '⊕' means physical mixture. The quantities $P_{S,1}$, $P_{S,2}$ and $P_{S,3}$ are called the tristimulus values. Their subscripts identify the test light (S) and the primaries (1, 2 or 3). Using Grassmann's law we treat this mixture equation as an algebraic equation and express the match in terms only of S:

$$P_S S = P_{S,1}\mathbf{P}_1 + P_{S,2}\mathbf{P}_2 - P_{S,3}\mathbf{P}_3 \qquad (3.2)$$

where the minus sign reflects the fact that in the color match, the primary $\mathbf{P}_3$ actually was added to the test light. Given a set of primaries, ($\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$), a color match can be made to all lights, of any spectral power distribution. When the test light is a narrow spectral band, one of the primaries is always either negative; i.e. physically superimposed with the spectral test light to match the remaining two primaries, or zero. When the test light has a broad spectral power distribution, the three primaries may all be positive.

As a consequence of Grassmann's law, the match to a broad spectral power distribution can be considered as the sum of constituent narrow-band spectral matches. Consider a distribution of unit energy at every spectral wavelength (the equal energy spectrum, abbreviated EES). According to the law of additivity, the color match to this distribution is considered as the sum of the matches to unit energies of the spectral wavelengths.

$$P_{EES}EES = [\Sigma(P_\lambda)]\ EES = [\Sigma(P_{\lambda,1})]\ \mathbf{P}_1 + [\Sigma(P_{\lambda,2})]\ \mathbf{P}_2 + [\Sigma(P_{\lambda,3})]\ \mathbf{P}_3 \qquad (3.3)$$

Data of color matching: Early measurements of color matching are reviewed by Bouma (1947) and Le Grand (1968). The data on which 2° colorimetry was formulated were those of Wright (1929, 1946) and Guild (1931). Stiles (1955) has discussed this work and presented modern data for both 2° and 10° fields. Here we describe Wright's experiment and how the colorimetric observer was formulated.

Wright used a technique known as maximum saturation matching. The primaries and the spectral test lights are arranged pairwise in a bipartite field so that a mixture of the spectral test light and one primary are matched to a mixture of the two other primaries. The spectral primaries were at 650 nm ($\mathbf{P}_1$), 530 nm ($\mathbf{P}_2$), and 460 nm ($\mathbf{P}_3$). The spectral test wavelengths varied in 10 or 20 nm steps between 410 and 700 nm. For each test wavelength, a color match was obtained and the amount of each primary was measured and expressed as an equation. Every spectral wavelength could be specified in terms of the three primaries, with one of the three being negative or zero. Wright also made a match to his instrument broadband (white) source, which was filtered tungsten with a color temperature of 4800 K. This source was essentially equivalent to one adopted formally by the CIE and termed Standard Illuminant B.[1]

3.2.2 REPRESENTATION OF COLOR MATCHING DATA

The quantities of the test wavelengths are customarily referred to an equal energy spectrum. There are several conventions to represent the unit of measurement for the primaries (Wright, 1946; Stiles, 1955; Stiles and Burch, 1959; Le Grand, 1968). It is convenient for the purpose of discussion to redefine the tristimulus values $P_{S,1}$, $P_{S,2}$, and $P_{S,3}$ of equations 3.1–3.2, using new terms, $C_{S,1}$, $C_{S,2}$, and $C_{S,3}$:

$$C_{S,i} = P_{S,i}/e_i \qquad (3.4)$$

where $P_{S,i}$ is the radiant flux of primary (i) specified relative to a radiant unit, e_i. If the three radiant units are equivalent and unity, the data will be expressed in watts, as implied in equations 3.1–3.3. In practice, however, other normalizations are used. In energy units, the short-wavelength $\mathbf{C}_1$ primary amount is much higher than the others. In representing color

mixture data, it was considered desirable that the amounts of the primaries have similar scales. Recalculating a new normalization simply requires weighting all the values of C_i. Suppose e'_i is the new primary unit:

$$P_{S,i} = C_{S,i}\, e'_i = C_{S,i}\, e_i \qquad (3.5)$$
$$C_{S,I} = C_{S,i}\, e_i / e'_i \qquad (3.6)$$

Chromaticity diagrams: The data of a color matching experiment can be expressed in terms of vectors in a three-dimensional space (Schrödinger, 1920). The tristimulus values, $C_{1,1}$, $C_{2,2}$, and $C_{3,3}$, at the three primaries form unit vectors. A resultant vector represents the location of a mixture of the three primaries. According to the Grassmann laws, the scaling of the unit vectors is arbitrary; if the radiance were doubled, the entire diagram would be expanded but would maintain its shape. A two-dimensional unit plane can be defined in the space where the sum of the tristimulus values is unity. The relative positions of the resultant vectors are preserved in such a plane. The tristimulus values are converted into a form where the sum of the three always equals unity. For a test color S:

$$c_{S,1} = C_{S,1}/[C_{S,1} + C_{S,2} + C_{S,3}], \qquad (3.7)$$
$$c_{S,2} = C_{S,2}/[C_{S,1} + C_{S,2} + C_{S,3}] \qquad (3.8)$$
$$c_{S,3} = C_{S,3}/[C_{S,1} + C_{S,2} + C_{S,3}]. \qquad (3.9)$$

The quantities $c_{S,i}$ are termed chromaticity coordinates (trichromatic coefficients in the older literature). The unit plane for Wright's data is plotted in Cartesian coordinates in Figure 3.1. The horseshoe shape shows the location of the spectral test wavelengths and the head of vector B represents the location of Wright's instrument white (Standard Illuminant B).

The chromaticity diagram is a two-dimensional representation of the results of a color mixture experiment. Figure 3.2 shows Wright's data replotted in a chromaticity diagram where c_1 is plotted against c_2 on Cartesian axes. The coordinates for $\mathbf{P}_1$, $\mathbf{P}_2$, and $\mathbf{P}_3$ form a right-angled triangle. Points that fall outside this triangle represent the negative values of the spectral tristimulus values. The values for the spectral wavelengths again demonstrate a characteristic horseshoe-shaped curve; this is called the spectrum locus. Standard Illuminant B appears near the center of the right-angled triangle as point S_B.

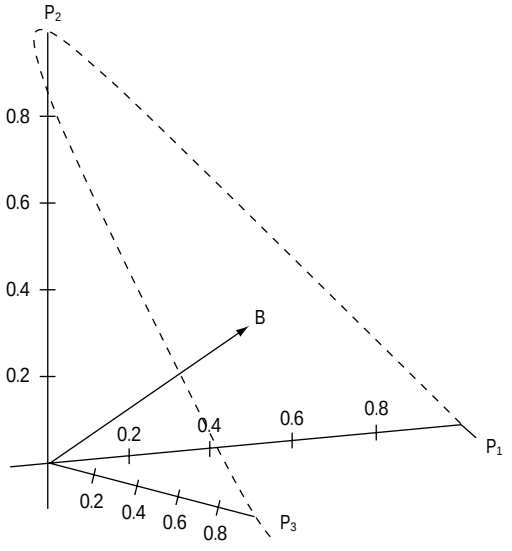


Figure 3.1 **Representation of Wright's color matching data in a three-dimensional space. The three primaries P_1, P_2, P_3 form the axes. Wright's chromaticity coordinates for the spectral test lights are shown plotted in the unit plane. The arrowhead represents the vector for source B.**

WDW normalization: W.D. Wright performed his experiment using a method of normalization now termed WDW normalization. This normalization is a practical choice, because it ensures that the tristimulus values will have rather similar weights, but does not require the experimenter to have an absolute calibration of the radiant energies of the primaries and test lights. There is additionally a theoretical significance to data expressed in this normalization.

In WDW normalization, two test wavelengths, the first intermediate between $\mathbf{P}_1$ and $\mathbf{P}_2$ and the second intermediate between $\mathbf{P}_2$ and $\mathbf{P}_3$, are chosen as normalizing wavelengths. Wright used 582.5 nm and 494 nm. The energy units are chosen so that the amount C_1 of the $\mathbf{P}_1$ primary is set equal to the amount C_2 of the $\mathbf{P}_2$ primary at the match to the first normalizing wavelength and the amount C_3 of the $\mathbf{P}_3$ primary is set equal to the amount C_2 of the $\mathbf{P}_2$ primary at the match

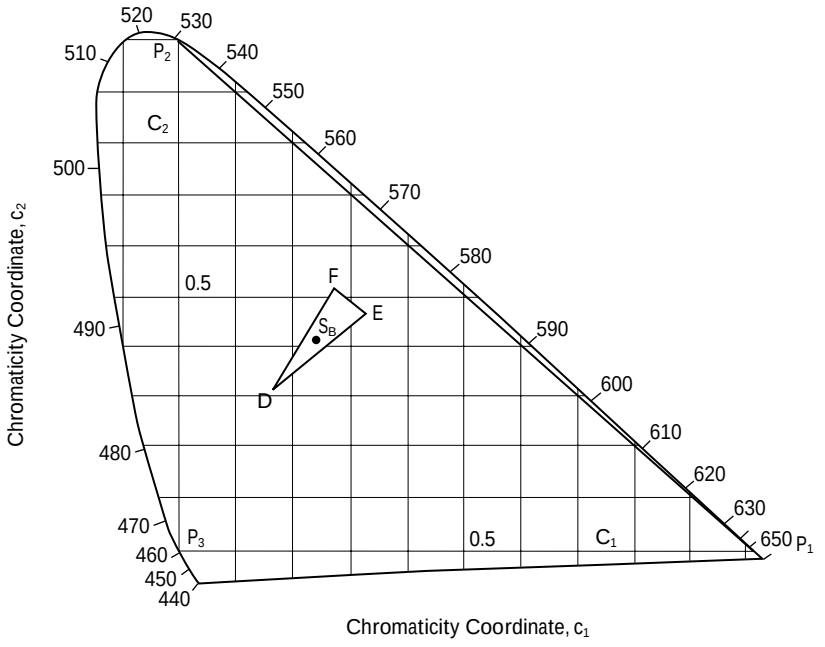


Figure 3.2 The chromaticity diagram calculated for Wright's data with WDW normalization and C_2 plotted versus C_1. The coordinates for S_B represent the position of the average match to Standard Illuminant B made by 36 observers, at coordinates (0.243, 0.410). The triangle DEF shows the spread of the individual white matches. (From Y. Le Grand, *Light, Colour, and Vision*, translated by R.W.G. Hunt, J.W.T. Walsh, and F.R.W. Hunt, Dover Publications, New York, 1957; reprinted with permission.)

to the second normalizing wavelength. In Figure 3.2, the spectrum locus is the averaged data of 10 observers. The coordinates for S_B represent the position of the average match to the instrument white made by 36 observers, at coordinates (0.243, 0.410). The triangle DEF shows the spread of the individual white matches. Figure 3.3 shows the chromaticity coordinates for the spectrum for Wright's observers plotted as a function of wavelength. The solid line shows the average data for 10 observers and the dashed lines show the matches of extreme observers. The interobserver variability is partitioned among the spectral wavelengths and the white point.

The importance of the WDW normalization is that it separates interobserver variance caused by receptor variation from interobserver variance caused by prereceptoral variation. It may be shown algebraically that in the WDW system, interobserver variation in the chromaticity coordinates for the spectrum greater than experimental error can be attributed to interobserver differences in the spectral absorption characteristics of the visual photoreceptors. Individual variation in the distribution of the coefficients for the white point greater than experimental error must be attributed to variation in prereceptoral filters, e.g. the lens and macular pigment (Wright, 1946; Wyszecki and Stiles, 1982) (see Chapter 2).

Photometry and colorimetry: The discussion until now has assumed that the primary units are specified in radiant energy. However, that is not how colorimetry developed. Historically, the radiant energy levels were not available in the early studies of Wright and Guild. Wright circumvented this problem by use of the WDW normalization and he expressed his data in chromaticity coordinates.

Equation (3.4) can also be expressed in luminous quantities, F:

$$F_{\lambda l} = K_m V_{\lambda l} P_\lambda \qquad (3.10)$$

where K_m is the constant relating lumens to watts (683 lumens/watt) and V_λ is the spectral

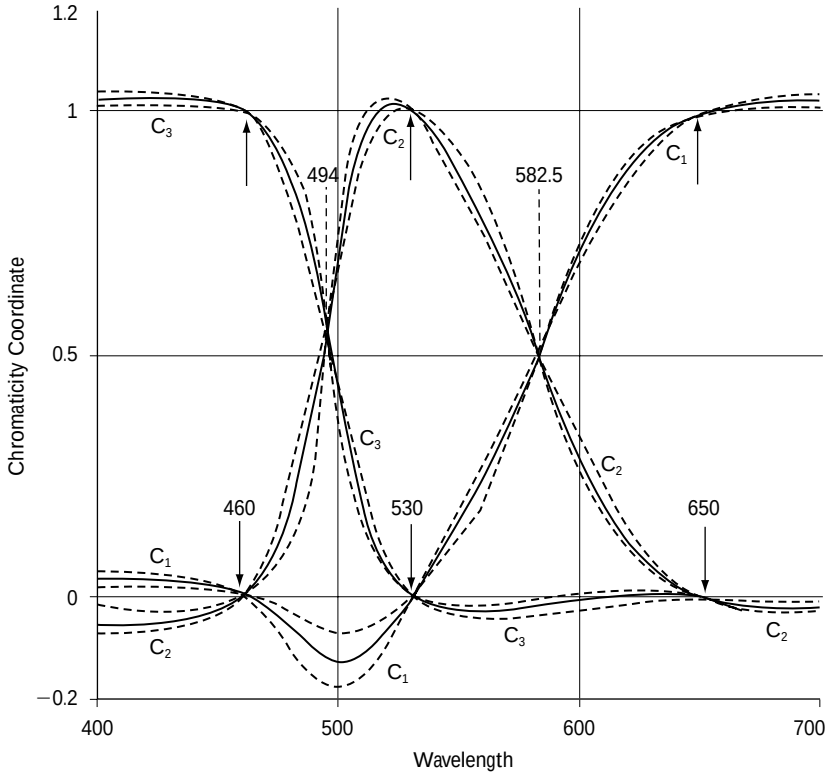


Figure 3.3 **Wright's chromaticity coordinates plotted as a function of wavelength with WDW normalization. The solid line represents the average of 10 observers; the dashed lines show the matches of extreme observers. (From Y. Le Grand, *Light, Colour and Vision*, translated by R. W. G. Hunt, J. W. T. Walsh, and F. R. W. Hunt, Dover Publications, New York, 1957; reprinted with permission.)**

luminous efficiency of the human eye. The V_λ had been adopted as a standard by the CIE in 1924 (CIE, 1926), just a few years before Wright's experiments. Equation (3.4) can thus be rewritten:

$$C_{\lambda,i} = F_{\lambda,i}/l_i \qquad (3.11)$$

where $F_{\lambda,i}$ is the luminous flux of primary light, (i) specified relative to its luminous unit, l_i. Wright measured the luminous units for the primaries using heterochromatic flicker photometry. The measured values were adjusted to predict the chromaticity coefficients of Standard Illuminant B, consistent with the V_λ curve. The relative luminance of the $\mathbf{P}_1$, $\mathbf{P}_2$, and $\mathbf{P}_3$ primaries was calculated to be in the ratio of 0.65:1:0.044.

The colorimetric match is a statement of identity of appearance. At the match, the luminous fluxes must also be equivalent. Therefore for any test light S, of luminous flux F_S:

$$F_S = F_{S1} + F_{S2} + F_{S3} = l_1C_{S1} + l_2C_{S2} + l_3C_{S3} \qquad (3.12)$$

For the spectral test lights referred to, an equal energy spectrum, F_λ is proportional to V_λ. Further, the luminous unit, l_λ for a spectral test light, C_λ of luminous flux F_λ can be defined:

$$l_\lambda = F_\lambda/C_\lambda = (l_1C_{\lambda,1} + l_2C_{\lambda,2} + l_3C_{\lambda,3})/(C_{\lambda,1} + C_{\lambda,2} + C_{\lambda,3}) \qquad (3.13)$$

$$= l_1c_{\lambda,1} + l_2c_{\lambda,2} + l_3c_{\lambda,3} \qquad (3.14)$$

If the spectral chromaticity coordinates and the luminous units, l_i, of the primaries are known, the spectral tristimulus values can be derived from the equations above:

$$C_{\lambda,i}(\lambda) = c_{\lambda,i}V(\lambda)/(l_1c_{\lambda,1} + l_2c_{\lambda,2} + l_3c_{\lambda,3}) \qquad (3.15)$$

The data for Figure 3.1 were derived in this manner from Wright's tabulation of the chromaticity coefficients and the values of the relative

luminous units given above. In modern terminology, for a set of three primaries, identified as $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$, the tristimulus values are identified as $P_{\lambda,1}$, $P_{\lambda,2}$, and $P_{\lambda,3}$, and the chromaticity coordinates are $p_{\lambda,1}$, $p_{\lambda,2}$, and $p_{\lambda,3}$. The tristimulus values for an arbitrary light, S, are $P_{S,1}$, $P_{S,2}$, and $P_{S,3}$.

Linear transformations of color matching data: Grassmann's (1853) observation that color mixture data show the associative property allows expression of a set of data in sets of primaries other than those used in a particular experiment. This can be noted intuitively since any test wavelength is a linear combination of a given primary set ($\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$). These primaries can themselves be considered as a linear sum of some other primary set ($\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{Q}_3$). The equations are more economically set up in matrix notation.

A transformation from one set of primaries ($\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$) to another set ($\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{Q}_3$) is specified by a general homogeneous linear transformation. The rules for choosing the new primaries ensure that the matrix A relating ($\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$) to ($\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{Q}_3$) has an inverse (i.e. the determinant of A ≠ zero). A test color in the original primary set can be described in terms of either primary set. In particular, a given test color S with tristimulus values $P_{S,i}$ in Primary set ($\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{Q}_3$):

$$\begin{bmatrix} P_{S,1} \\ P_{S,2} \\ P_{S,3} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_{S,1} \\ \mathbf{Q}_{S,2} \\ \mathbf{Q}_{S,3} \end{bmatrix} \qquad (3.16)$$

has tristimulus values $Q_{S,i}$ in Primary set ($\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$):

$$\begin{bmatrix} Q_{S,1} \\ Q_{S,2} \\ Q_{S,3} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{S,1} \\ \mathbf{P}_{S,2} \\ \mathbf{P}_{S,3} \end{bmatrix} \qquad (3.17)$$

where matrix B is the inverse of matrix A. In particular, if the tristimulus values at the new primary set ($\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{Q}_3$) are known in the original set ($\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$), the matrix A can be written explicitly:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} {}^{P}\mathbf{Q}_{1,1} & {}^{P}\mathbf{Q}_{2,1} & {}^{P}\mathbf{Q}_{3,1} \\ {}^{P}\mathbf{Q}_{1,2} & {}^{P}\mathbf{Q}_{2,2} & {}^{P}\mathbf{Q}_{3,2} \\ {}^{P}\mathbf{Q}_{1,3} & {}^{P}\mathbf{Q}_{2,3} & {}^{P}\mathbf{Q}_{3,3} \end{bmatrix} \qquad (3.18)$$

where ${}^{P}\mathbf{Q}_{1,1}$, ${}^{P}\mathbf{Q}_{1,2}$, and ${}^{P}\mathbf{Q}_{1,3}$ are the tristimulus values of primary $\mathbf{Q}_1$, ${}^{P}\mathbf{Q}_{2,1}$, ${}^{P}\mathbf{Q}_{2,2}$, and ${}^{P}\mathbf{Q}_{2,3}$ are the tristimulus values of primary $\mathbf{Q}_2$, and ${}^{P}\mathbf{Q}_{3,1}$, ${}^{P}\mathbf{Q}_{3,2}$, and ${}^{P}\mathbf{Q}_{3,3}$ are the tristimulus values of primary $\mathbf{Q}_3$ in the $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$ primary set. These assignments uniquely determine the transformation. However, the data will be normalized to the energy units of the spectral matches. To obtain the same normalization as the original $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$ primary set, a diagonal matrix containing the inverse of the energy units, e_i is used to multiply matrix B.

The properties of a homogeneous linear transformation (also called an affine transformation) include: straight lines remain straight after transformation; parallel lines remain parallel; and plane surfaces remain plane surfaces.

Transformations may also be made directly between chromaticity coefficients. If matrix B is known, the chromaticity coordinates, q_1 and q_2, are given by:

$$q_1 = b11p_1 + b12p_2 + b13p_3/[(b11 + b12 + b13)p_1 + (b12 + b22 + b23)p_2 + (b31 + b32 + b33)p_3] \qquad (3.19)$$

$$q_2 = b21p_1 + b22p_2 + b23p_3/[(b11 + b12 + b13)p_1 + (b12 + b22 + b23)p_2 + (b31 + b32 + b33)p_3] \qquad (3.20)$$

The transformation of chromaticity coefficients is projective rather than linear. Straight lines remain straight, but parallel lines need not remain parallel.

The subset of physically realizable lights may be thought of as only a small portion of a chromaticity space. The possibility exists of defining new primaries determined by locations in chromaticity space that lie outside the spectrum locus. In this case, the tristimulus values are not known. A transformation matrix, B′ is determined, except for an arbitrary scaling factor, by four unique loci in the chromaticity space, the three new primary coordinates and a fourth locus, usually the equal energy spectrum. These coordinates provide coefficients for three sets of simultaneous equations, one for each new primary, which are solved to yield the elements of matrix B′.

Primary transformations have been used for widely diverse purposes, including: comparison

of data sets; derivation of an all-positive (imaginary) primary system (Judd, 1930); derivation of color planes in which equal vector lengths represent equal steps in discriminability (Judd, 1935; Galbraith and Marshall, 1985); derivation of a coordinate system in which the primaries represent the cone spectral sensitivities (König and Dieterici, 1893; Vos and Walraven, 1971; Smith and Pokorny, 1975; Stockman *et al.*, 1993; Stockman and Sharpe, 2000); and derivation of a coordinate system in which the primaries represent physiologically relevant opponent channel sensitivities (Schrödinger, 1925; Judd, 1951a).

3.2.3 CIE STANDARD COLORIMETRIC OBSERVERS

The industrial importance of trichromacy is in prediction of visual equivalency of a wide array of spectral power distributions. The colorimetric data provide a numerical specification at unit wavelength steps from which can be calculated tristimulus values for any spectral power distribution. The CIE in 1931 defined a standard observer for colorimetry, based on 2° color matching. The 2° observer is recommended for fields up to 4°. A 10° observer was defined in 1964 (CIE, 1964). The characteristics of the large-field observer are recommended for visual stimuli whose extent exceeds 4, but should only be used at high photopic illuminances. Revision and evaluation of the standard observers remains a current interest of the CIE.

The 2° observer: The 1931 CIE standard observer for colorimetry incorporates both colorimetric and photometric behavior. The basic data were averaged chromaticity coefficients of Wright (1929) and Guild (1931). These were expressed in Wright's primaries and normalized to Standard Illuminant B. The luminous units of the primaries were adjusted to be consistent with the location of Standard Illuminant B and with the luminosity of the 1924 CIE standard observer for photometry.

Two equivalent statements of the color-matching behavior of the 1931 CIE standard observer were embodied in the (**R,G,B**) and (**X,Y,Z**) systems of units. The (**R,G,B**) system represented a first step in which spectral primaries were retained but all the negative values were in the $G(\lambda)$. Normalization was to the equal energy spectrum. An important consideration was that, following transform of the chromaticity coordinates, color matching functions using equation (3.15) were determined by incorporating $V(\lambda)$, the luminous efficiency function of the 1924 Standard observer for photometry. The 1924 CIE photometric observer was based on an entirely different set of observations than the color matching. The accuracy of the 1931 CIE colorimetric observer is thus dependent both on the accuracy of the Wright and Guild color matching data and on the accuracy of the 1924 CIE spectral luminous efficiency function. Stiles presented color matching data based on absolute radiometric calibration. His calculated chromaticity coefficients were very close to those of Wright and Guild.

The (**X,Y,Z**) system is an all-positive system derived by a linear transformation of the (**R,G,B**) system. An all-positive system means that the primaries enclose the spectrum locus; the primaries are not realizable lights and are sometimes called 'imaginary' primaries. The transformation was defined so that the color matching function, Y_λ was equivalent to $V(\lambda)$, the CIE photopic luminous efficiency function for the standard observer. Therefore, the **X** and **Z** primaries have zero luminance. In color space the line of zero luminance is called the alychne (Schrödinger, 1920). Given that the luminous units of the (**R,G,B**) system are in the ratio 1:4.5907:0.0601, the equation for the alychne obeys:

$$r + 4.5901g + 0.0601b = 0 \qquad (3.21)$$

The **X** and **Y** primaries were placed on a line connecting the color matches made above 550 nm. This means that **Z** has no contribution to color matching at long wavelengths. The **Y** and **Z** primaries were placed on a line tangent to the spectrum locus at 504 nm. This choice was to minimize the area between the unit triangle and the spectrum locus. The intersections of the three lines determine the chromaticity coordinates (r, g) of the new primaries. Normalization was to the equal-energy white. The (**X,Y,Z**) system is specified by a set of color matching functions, X_λ, Y_λ, and Z_λ shown in Figure 3.4 and a chromaticity diagram, x_λ, y_λ shown in Figure 3.5.

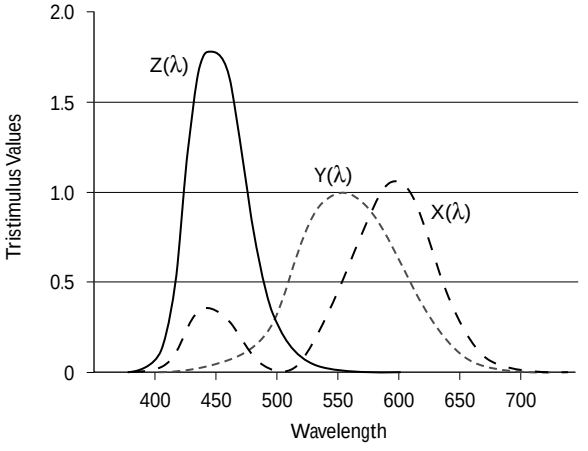


Figure 3.4 **CIE 1931 color matching functions. Spectral tristimulus values of constant radiance stimuli for different wavelengths. The three functions of wavelength define the color-matching properties of the CIE 1931 standard colorimetric observer. (Plotted from Table 3.1.)**

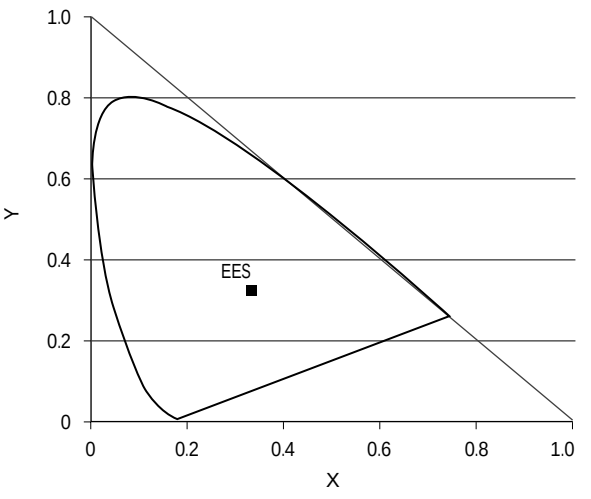


Figure 3.5 **CIE 1931 chromaticity diagram. CIE 1931 (x,y)-chromaticity diagram. (Plotted from Table 3.1.)**

The (**X,Y,Z**) system is widely used for color - specification in industry. Table 3.1 shows color matching functions and chromaticity coordinates at 10 nm intervals.

The Judd (1951) modified 2° observer (Judd 1951b): After the acceptance of the 1931 CIE observer, problems were noted in the estimates of continuous spectral power distributions. The difficulties were ascribed to underestimation of luminosity at short wavelengths in the 1924 CIE photometric observer. Judd proposed a revision of the 1931 CIE colorimetric observer, with a modification of the Y_λ function below 460 nm and a slight modification of the chromaticity coordinates. The Judd revised colorimetric observer gave new functions, $X_{J,\lambda}$, $Y_{J,\lambda}$, and $Z_{J,\lambda}$ in the short wavelength region. The Judd revised colorimetric observer did not replace the CIE 2° observer, which remains widely used in industry. However, the Judd revised colorimetric observer is frequently used in color vision theory. Smith, Pokorny, and Zaidi (1983) observed that the Judd revised colorimetric observer is characterized by less lens and more prereceptoral macular pigment absorption than is the CIE 2° observer. Table 3.2 shows color matching functions and chromaticity coordinates at 10 nm intervals for the Judd (1951) revised observer.

In 1988, the CIE recommended a supplementary observer for photometry, termed $V_M(\lambda)$. This observer differs from the Judd revised $y_{J\lambda}$ below 410 nm. The further revision to give $V_M(\lambda)$ was originally proposed by Vos (1978), who also provided tabulations for an amended colorimetric observer.

The 10° observer: The 1964 CIE large-field standard observer for colorimetry was based on color matching data from the laboratories of Stiles and Burch (1959) and Speranskaya (1959, 1961). Since cone-dependent color matching functions were desired, the experimenters were concerned to avoid rod intrusion (changes in color matches caused by rod activity). Stiles and Burch (1959) maintained high photopic luminance of the matching fields. Multiple primary sets were used to minimize rod contamination of the color matches. Finally, Stiles and Burch (1959) made mathematical corrections to the data using a theoretical expectation of the nature of rod intrusion (see also Wyszecki and Stiles, 1982; Shapiro *et al.*, 1994).

The CIE transformed the data into an all-positive system with properties similar to those of the 1931 (**X,Y,Z**) system. The 1964 10° standard observer is specified by a set of color matching functions, $X_{10}(\lambda)$ $Y_{10}(\lambda)$, and $Z_{10}(\lambda)$ and a chromaticity diagram, $x_{10}(\lambda)$, $y_{10}(\lambda)$. The $Y_{10}(\lambda)$ represents the relative spectral luminous

Table 3.1 Color matching functions and chromaticity coordinates for the CIE (1931) Standard colorimetric observer

Wavelength	X	Y	Z	x	y
380	0.0014	0	0.0065	0.1772	0.0000
390	0.0042	0.0001	0.0201	0.1721	0.0041
400	0.0143	0.0004	0.0679	0.1731	0.0048
410	0.0435	0.0012	0.2074	0.1726	0.0048
420	0.1344	0.0040	0.6456	0.1714	0.0051
430	0.2839	0.0116	1.3856	0.1689	0.0069
440	0.3483	0.0230	1.7471	0.1644	0.0109
450	0.3362	0.0380	1.7721	0.1566	0.0177
460	0.2908	0.0600	1.6692	0.1440	0.0297
470	0.1954	0.0910	1.2876	0.1241	0.0578
480	0.0956	0.1390	0.813	0.0913	0.1327
490	0.032	0.2080	0.4652	0.0454	0.2950
500	0.0049	0.3230	0.272	0.0082	0.5384
510	0.0093	0.5030	0.1582	0.0139	0.7502
520	0.0633	0.7100	0.0782	0.0743	0.8338
530	0.1655	0.8620	0.0422	0.1547	0.8058
540	0.2904	0.9540	0.0203	0.2296	0.7543
550	0.4334	0.9950	0.0087	0.3016	0.6924
560	0.5945	0.9950	0.0039	0.3731	0.6245
570	0.7621	0.9520	0.0021	0.4441	0.5547
580	0.9163	0.8700	0.0017	0.5125	0.4866
590	1.0263	0.7570	0.0011	0.5752	0.4242
600	1.0622	0.6310	0.0008	0.6270	0.3725
610	1.0026	0.5030	0.0003	0.6658	0.3340
620	0.8544	0.3810	0.0002	0.6915	0.3084
630	0.6424	0.2650	0	0.7080	0.2920
640	0.4479	0.1750	0	0.7191	0.2809
650	0.2835	0.1070	0	0.7260	0.2740
660	0.1649	0.0610	0	0.7300	0.2700
670	0.0874	0.0320	0	0.7320	0.2680
680	0.0468	0.0170	0	0.7335	0.2665
690	0.0227	0.0082	0	0.7346	0.2654
700	0.0114	0.0041	0	0.7347	0.2653
710	0.0058	0.0021	0	0.7347	0.2653
720	0.0029	0.0010	0	0.7347	0.2653
730	0.0014	0.0005	0	0.7347	0.2653
740	0.0007	0.0003	0	0.7347	0.2653
750	0.0003	0.0001	0	0.7347	0.2653
760	0.0002	0.0001	0	0.7347	0.2653
770	0.0001	0	0	0.7347	0.2653
780	0	0	0	0.7347	0.2653

efficiency function of the 10° standard observer, although it has never been accepted as a standard for photometry. Since the data are representative of matches made by an observer without rod function, 10° matches made by color normal observers with other primaries and luminances need not in general correspond to the 10° CIE Standard observer.

Properties of the chromaticity diagram: In the all-positive (**X,Y,Z**) system, an isosceles right triangle completely encloses the experimentally determined chromaticity diagram (Figure 3.5). The abscissa (y = 0) is the alychne. The spectrum locus forms its characteristic horseshoe shape. A line connects the coordinates for 380 nm and 700 nm and represents mixtures formed by the extreme short wavelength and the extreme long wavelength lights.

Chromaticity coordinates (x, y) may be calculated for light of any spectral power distribution. The tristimulus values of Q are given by:

Table 3.2 Color matching functions and chromaticity coordinates for the Judd (1951) Revised colorimetric observer. Values are tabulated at 10 nm intervals between 380 nm and 780 nm

Wavelength	X	Y	Z	x	y
380	0.0045	0.0004	0.0224	0.1648	0.0147
390	0.0201	0.0015	0.0925	0.1762	0.0131
400	0.0611	0.0045	0.2799	0.1768	0.0130
410	0.1267	0.0093	0.5835	0.1761	0.0129
420	0.2285	0.0175	1.0622	0.1747	0.0134
430	0.3081	0.0273	1.4526	0.1723	0.0153
440	0.3312	0.0379	1.6064	0.1677	0.0192
450	0.2888	0.0468	1.4717	0.1598	0.0259
460	0.2323	0.0600	1.2880	0.1470	0.0380
470	0.1745	0.0910	1.1133	0.1266	0.0660
480	0.092	0.1390	0.7552	0.0933	0.1409
490	0.0318	0.2080	0.4461	0.0464	0.3033
500	0.0048	0.3230	0.2644	0.0081	0.5454
510	0.0093	0.5030	0.1541	0.0140	0.7548
520	0.0636	0.7100	0.0763	0.0748	0.8354
530	0.1668	0.8620	0.0412	0.1559	0.8056
540	0.2926	0.9540	0.0200	0.2310	0.7532
550	0.4364	0.9950	0.0088	0.3030	0.6909
560	0.597	0.9950	0.0039	0.3741	0.6235
570	0.7642	0.9520	0.0020	0.4448	0.5541
580	0.9159	0.8700	0.0016	0.5124	0.4867
590	1.0225	0.7570	0.0011	0.5742	0.4251
600	1.0544	0.6310	0.0007	0.6253	0.3742
610	0.9922	0.5030	0.0003	0.6635	0.3363
620	0.8432	0.3810	0.0002	0.6887	0.3112
630	0.6327	0.2650	0.0001	0.7047	0.2952
640	0.4404	0.1750	0	0.7156	0.2844
650	0.2787	0.1070	0	0.7226	0.2774
660	0.1619	0.0610	0	0.7263	0.2737
670	0.0858	0.0320	0	0.7284	0.2716
680	0.0459	0.0170	0	0.7297	0.2703
690	0.0222	0.0082	0	0.7303	0.2697
700	0.0114	0.0041	0	0.7347	0.2653
710	0.0058	0.0021	0	0.7347	0.2653
720	0.0029	0.0010	0	0.7347	0.2653
730	0.0014	0.0005	0	0.7347	0.2653
740	0.0007	0.0003	0	0.7347	0.2653
750	0.0003	0.0001	0	0.7347	0.2653
760	0.0002	0.0001	0	0.7347	0.2653
770	0.0001	0	0	0.7347	0.2653
780	0	0	0	0.7347	0.2653

$$\begin{bmatrix} X_Q \\ Y_Q \\ Z_Q \end{bmatrix} = \begin{bmatrix} K \Sigma P(\lambda) \\ K \Sigma P(\lambda) \\ K \Sigma P(\lambda) \end{bmatrix} \begin{bmatrix} X(\lambda), \\ Y(\lambda), \\ Z(\lambda), \end{bmatrix} \quad (3.22)$$

$$= \quad (3.23)$$

$$= \quad (3.24)$$

where K is a constant and $P(\lambda)$ is the spectral power distribution. The chromaticity coordinates of Q are given by:

$$x_Q = X_Q/(X_Q + Y_Q + Z_Q), \quad (3.25)$$

$$y_Q = Y_Q/(X_Q + Y_Q + Z_Q). \quad (3.26)$$

Useful relations to transfer among the tristimulus values and the chromaticity coordinates are:

$$X_Q = (x_Q/y_Q)Y_Q \quad (3.27)$$

$$Z_Q = (z_Q/y_Q)\,Y_Q = ((1 - x_Q - y_Q)/y_Q)Y_Q \quad (3.28)$$

In equations 3.22–3.24, K is a normalizing constant that disappears from the chromaticity coordinates. In the definition of the 1924 CIE standard observer for photometry, K is identified

with k_m the conversion factor for lumens/watt. The value Y_Q can thus be expressed in lumens. The spectral power distribution Q is completely specified colorimetrically by its tristimulus values, X_Q, Y_Q, Z_Q, or alternatively by its chromaticity coordinates and luminance, x_Q, y_Q, Y_Q.

Representation of papers or filters in the x,y chromaticity diagram: There is an alternative way to consider K that is useful for specification of transmitting filters or reflective surfaces, such as papers. In this case, the colorimetric properties of the filter or paper under the illuminant are important but the absolute radiance level of the source is irrelevant. Suppose there is a pigment sample of spectral reflectance $\rho(\lambda)$ or a filter of spectral transmission $\tau(\lambda)$ viewed under an illumination, H of spectral distribution $H(\lambda)$. Equations 3.22–3.24 are rewritten to incorporate the spectral properties of the sample, S. For a pigment sample:

$$X_S = K \sum_{\lambda} \rho(\lambda)H(\lambda)\,X(\lambda), \qquad (3.29)$$

$$Y_S = K \sum_{\lambda} \rho(\lambda)H(\lambda)\,Y(\lambda), \qquad (3.30)$$

$$Z_S = K \sum_{\lambda} \rho(\lambda)H(\lambda)\,Z(\lambda). \qquad (3.31)$$

With K set at $100/\sum_{\lambda} H(\lambda)Y(\lambda)$, Y_S has the value of $100[\sum_{\lambda}\rho(\lambda)H(\lambda)Y(\lambda)]/[\sum_{\lambda}(\lambda)H(\lambda)Y(\lambda)]$. If the object is a perfectly diffusing surface, with $\rho(\lambda) = 1.0$ for all wavelengths, or a perfectly transmitting object, with $\tau(\lambda) = 1.0$ for all wavelengths, Y has the value 100. Thus, Y may be interpreted as the percentage luminous reflectance or percentage luminous transmittance of the sample. Calculation of the percentage luminous reflectance or transmittance is ratified only for the 1931 CIE colorimetric standard observer (Wyszecki and Stiles, 1982). Figure 3.6 shows the necessary calculations for a Wratten gelatin (No. 78) and CIE Standard Illuminant A[1].

A sample whose chromaticity coordinates are known may be specified by its dominant wavelength, λ_d, and its excitation purity, p_e, (Figure 3.6) for a specified CIE Illuminant. The dominant wavelength of sample S for Standard Illuminant A is the wavelength occurring at the intersection of the spectrum locus and a line extending from the locus of A through Q. If Q occurs in a region of the diagram for which the line extending from A to Q has no intersection with the spectrum locus, then the complementary wavelength may be used and should be noted by $-\lambda_d$ or λ_c. The excitation purity is the ratio of the distance from A to Q and the distance from A to λ_d:

$$p_e = (x_Q - x_A)/(x_d - x_A) = (y_Q - y_A)/(y_d - y_A) \quad (3.32)$$

In the case of a light whose dominant wavelength is $-\lambda_d$, the intersection with the line joining 380 nm and 700 nm is used. Excitation purity is zero when A and Q coincide and unity when Q and λ_d coincide. Both dominant wavelength and excitation purity may be estimated graphically. Precise computational methods for calculation of dominant wavelength are detailed by Wyszecki and Stiles (1982). The specification of dominant wavelength and excitation purity is widely used by manufacturers of colored filters.

Primary transformations using the CIE observers: Primary transformations using one of the CIE observers are simple because the tristimulus values (X,Y,Z) may be calculated for any choice of three lights. As an example of a problem applicable to visual science, consider the use of a color monitor system. The primaries are the three phosphors. All lights that can be produced on the monitor are combinations of possible light outputs of the three phosphors. Calibration of the relative spectral power distribution will yield the chromaticity coordinates (x_i, y_i) and calibration of the maximal luminance will yield $Y_{max,i}$ for phosphor (i). The luminance for phosphor (i) to light S is described by

$$Y_{S,i} = p_{S,i}\, Y_{max,i} \qquad (3.33)$$

where $p_{S,i}$ is the proportion of maximal phosphor output for phosphor (i). Further, the CIE tristimulus values for light S can be calculated from $Y_{S,1}$, $Y_{S,2}$, and $Y_{S,3}$:

$$\begin{aligned} X_S &= X_{S,1} + X_{S,2} + X_{S,3} \\ &= (x_1/y_1)Y_{S,1} + (x_2/y_2)Y_{S,2} \\ &\quad + (x_3/y_3)Y_{S,3} \end{aligned} \qquad (3.34)$$

$$Y_S = Y_{S,1} + Y_{S,2} + Y_{S,3} \qquad (3.35)$$

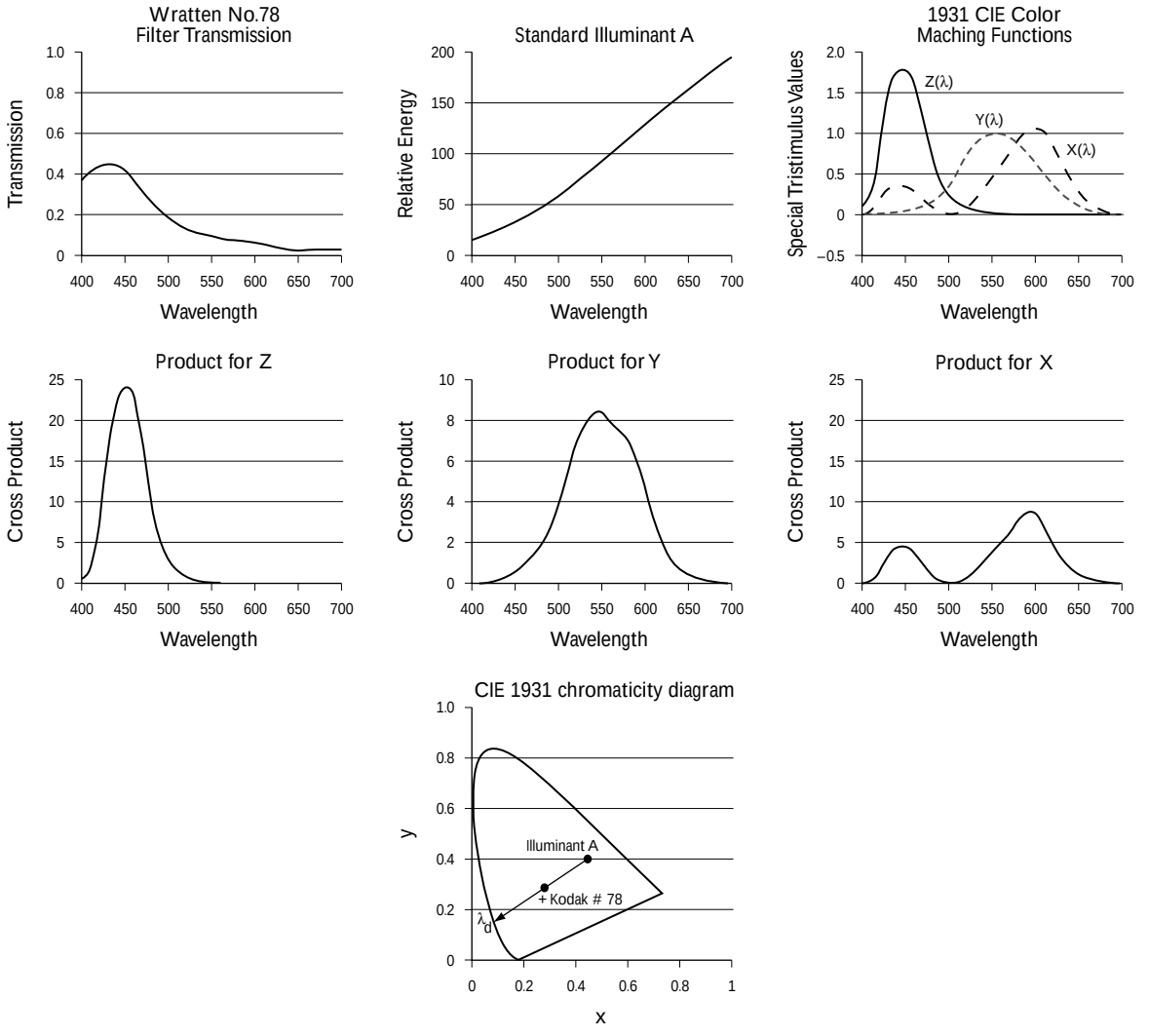


Figure 3.6 Calculation of tristimulus values and chromaticity coordinates for Wratten filter (No. 78) viewed under CIE Standard Illuminant A. The upper three panels show the transmission for the filter, the relative energy output for Standard Illuminant A and the 1931 CIE color matching functions. The middle three panels show the cross products of the filter transmission, CIE Standard Illuminant A, and the color matching functions. The lower panel shows the chromaticity diagram with the chromaticity coordinates of CIE standard illuminant A and the filter-illuminant combination. The line extending from CIE standard illuminant A, through the chromaticity coordinates of the filter-illuminant combination to the spectrum locus gives the dominant wavelength of 482 nm. The excitation purity is 0.45 and the luminous transmittance is 8.8%.

$$\begin{aligned} Z_S &= Z_{S,1} + Z_{S,2} + Z_{S,3} \\ &= (z_1/y_1)Y_{S,1} + (z_2/y_2)Y_{S,2} \\ &\quad + (z_3/y_3)Y_{S,3} \end{aligned} \tag{3.36}$$

This can be written in matrix form as:

$$\begin{bmatrix} X_S \\ Y_S \\ Z_S \end{bmatrix} = \begin{bmatrix} x_1/y_1 & x_2/y_2 & x_3/y_3 \\ 1 & 1 & 1 \\ z_1/y_1 & z_2/y_2 & z_3/y_3 \end{bmatrix} \begin{bmatrix} Y_{S,1} \\ Y_{S,2} \\ Y_{S,3} \end{bmatrix} \tag{3.37}$$

where $Y_{S,i}$ is given by equation (3.33). It is seen that matrix A contains only the chromaticity coefficients which are a characteristic of the phosphors. The inverse matrix gives the luminances of the phosphors required to produce a desired chromaticity and luminance specified in CIE coordinates.

3.2.4 EXPERIMENTAL VARIABLES

The Wright (1929) and Guild (1931) data were based on a 2° field viewed with foveal fixation. The method was maximum saturation matching in which the test light and one primary were matched to the remaining primaries in a bipartite field. The effects of procedure, field size, retinal illuminance, and fixation are important in color matching experiments and are detailed below.

Maxwell matching: In Maxwell matching, the spectral light plus a broadband light (such as the EES) are presented in one hemifield and compared to the three primaries presented in the other hemifield (Crawford, 1965). There are subtle differences in the color matching functions derived from the two methodologies which have eluded explanation (Zaidi, 1986). Stiles and Burch (1955) observed that the precision of color matches in the short wavelength region of the spectrum was improved by superimposing a mid-spectrum light on both fields.

Effect of field size: Color matches depend on the field of view; for example, color matching functions based on a 1.5° field (Fridrikh, 1957) show systematic differences from the 2° data (Pokorny *et al.*, 1976). A match for a 2° field will not hold for a larger or smaller field. There is a continuous change in the amounts of the primaries required for the color match as field size is increased (Pokorny and Smith, 1976).

For certain color stimuli in a 10° colorimetric field, a color inhomogeneity may appear at the area of fixation. This appears as an ill-defined ellipse with major axis horizontal, extending 1° or 2°, and is called the Maxwell spot. The spot follows fixation and is best observed by switching rapidly from viewing one half of the bipartite field to the other. In 10° trichromatic matching, the observer must ignore the Maxwell spot (e.g. Stiles and Burch, 1959). Alternatively an annular field may be used (e.g. Speranskaya, 1959, 1961). The Maxwell spot is usually attributed to higher density of macular pigment in the central 1–2° of the fovea. Trezona (1970) advanced an alternate hypothesis that the Maxwell spot might be a rod contrast color. This conclusion was based on her studies of 10° color matching using a tetrachromatic technique. A fourth primary was added to the mixture field to allow the observer to balance the rods in a dark-adapted state. By repeating the matching procedure in light- and dark-adapted states, a match is eventually obtained that holds in both light- and dark-adapted states. This 'tetrachromatic' match does not show a Maxwell spot. Palmer (1981) analyzed Trezona's paradigm and showed that her stimulus situation also reduced the differential effect of macular pigment in the two halves of the bipartite field.

When the color mixture field is reduced below 30′ of arc, there is a severe loss of discrimination. With stable fixation, if the field subtends 15′ or 20′ and is viewed either foveally or at 20′ or 40′ from the fovea, the normal trichromatic observer becomes dichromatic, requiring only two primaries for full spectrum color matching (Thomson and Wright, 1947). The color matches resemble those of the stationary congenital color defect, tritanopia (see section 3.5).

Effect of retinal illuminance: The scalar property states that metamers hold for all levels of retinal illumination; however, the range for which the 2° trichromatic metamers hold is limited to about 1–8000 td. Chromatic discrimination is optimal in a similar range of retinal illuminance. When the retinal illuminance exceeds 5000–10 000 td, lights that were metamers are no longer perceived as such (Brindley, 1953; Terstiege, 1967; Alpern, 1979; Wyszecki and Stiles, 1980). For example, in the match of 589 nm to a mixture of 545 nm and 670 nm primaries, a greater proportion of the 670 nm primary is required as retinal illuminance increases above 8000 td. The retinal illuminance at which the metamers break down is a level at which photopigment is significantly depleted by bleaching. While the change in match is a failure of the scalar property, the matches remain trichromatic. The effect and its explanation are described further in section 3.2.5.

With reduction in retinal illuminance, color matches continue to hold as low as 1 td. One effect of reduction in retinal illuminance is discrimination loss, similar to small-field tritanopia (Grigorovici and Aricescu-Savopol, 1958). With further reduction in luminance, rod intrusion becomes evident (Richards and Luria, 1964). Whether this is due to rod function in a foveal 2°

area or to changes in fixation with decreases in luminance is unknown. Estévez (1979) suggested that the Wright (1929), Guild (1931), and Stiles (1955) 2° data have a small rod contribution. Pokorny and Smith (1981) and Pokorny, Smith, and Went (1981) noted rod contribution in some color-defective individuals for 50 td fields as small as 1°. Rod intrusion is more easily seen in the color-defective observer whose cone-dominated color vision is compromised.

Peripheral color matching: Color matching may also be performed using the parafoveal or peripheral retina. With the dark-adapted eye and a scotopic illuminance, the normal observer is monochromatic over most of the spectrum, using the rod mechanism for color matching. At mesopic and photopic levels, parafoveal and peripheral color matching is trichromatic. Color matching data that allow the description of peripheral color matches in terms of foveal color vision have been described by Moreland and Cruz (1959) using an asymmetric matching procedure. For a spectral test field of $40' \times 80'$ viewed by the periphery and a $40' \times 80'$ mixture field viewed by the fovea with WDW normalization (section 3.2.2 above) at the foveal match, Moreland and Cruz determined the proportions of foveal primaries necessary to match various peripherally viewed stimuli. The chromaticity coefficients were all inside rather than on the spectrum locus for Wright's 2° foveal data, indicating that a spectral radiation of fixed size viewed in the peripheral retina is desaturated in appearance compared with its appearance at the fovea. Abramov, Gordon, and Chan (1991) have emphasized that a peripheral stimulus must be larger than a foveal stimulus in order to appear colored (see sec 4.2.1 in Chapter 4). Stabell and Stabell (1976) used the Moreland and Cruz paradigm to evaluate rod contribution to color matching by taking measurements during the cone-plateau period following intense light adaptation and noted that rod participation produces changes in all three chromaticity coordinates.

3.2.5 INTERPRETATION OF COLOR MATCHING

It was acknowledged by Young (1802) that every spectral wavelength is not recognized independently in the visual system. He proposed three physiological mechanisms or fundamentals that are differentially sensitive in the visible spectrum. In principal, the limitation of trichromacy could occur at any stage in the visual system (Brindley, 1970). Brindley proposed that foveal trichromacy is photopigment-limited. However, with a larger matching field or with parafoveal viewing, a fourth photoreceptor type, the rod, becomes active. Color matching remains trichromatic but does not obey Grassmann's laws. In this case, trichromatic color matching is neurally limited (Smith and Pokorny, 1977).

The quantal hypothesis: Brindley formalized the biological interpretation of color matching with his proposal that 2° foveal color matching is achieved when the quantal catch rate is equivalent for each active photopigment (Brindley, 1970). Since three sets of cone photopigments are active in the normal 2° fovea, color matching is trichromatic and photopigment-limited. Thus, foveal color matching can be considered a powerful tool in understanding human color vision, since the appropriate linear transform will reveal the effective absorption spectra of the cone visual photopigments. At the same time, color matching is a limited tool in understanding human color vision since it can give us no information on the subsequent neural transformations performed by the retina and visual cortex.

The spectral sensitivity of the fundamentals must be a linear transform of the color matching functions. However, since there are infinitely many possible linear transformations of color matching functions, external criteria are needed to guide and limit the transformation. The classic assumption, made by König was that congenital color-defectives (section 3.5) represented a reduction form of normal color vision. An early set of fundamentals, which relied on color matching data from congenital color-defective observers was published by König and Dieterici (1886, 1893). Fundamentals that rely on properties of color-defective vision are called König fundamentals.

Cone fundamentals: König hypothesized that the dichromatic forms of color defect represent reduced forms of normal color vision. Modern fundamentals are based on the König hypothesis

(Vos and Walraven, 1971; Smith and Pokorny, 1975; Stockman *et al.*, 1993; Stockman and Sharpe, 2000). We term the fundamentals S_λ, M_λ, and L_λ to indicate that these fundamentals explicitly represent the energy-based, corneal spectral sensitivities of the short-wavelength (SWS), middle-wavelength (MWS), and long-wavelength (LWS) sensitive cones respectively.

Vos and Walraven (1971) suggested that the Judd (1951) revised observer should be used to derive fundamentals. They further postulated that the three fundamentals should add to make the Judd (1951) luminous efficiency function, $y_{J,\lambda}$. Smith and Pokorny (1975) suggested that the luminous efficiency function should be determined only by $L(\lambda)$ and $M(\lambda)$, giving the property that the relative height of $S(\lambda)$ is undetermined. Their values for $L(\lambda)$ and $M(\lambda)$ differ only at short wavelengths from Vos and Walraven. Since the protanopic and deuteranopic copunctal points fall on the inverse diagonal of the spectrum locus in the chromaticity diagram, both transformations equate $S(\lambda)$ with $Z(\lambda)$. The Smith and Pokorny (1975) transformation equations, with $S(\lambda)/Y_j(\lambda)$ scaled to have a height of unity at 400 nm, are:[2]

$$L_\lambda = 0.15516\, X_{J,\lambda} + 0.54307\, Y_{J,\lambda}\, 0.03287\, Z_{J,\lambda} \quad (3.38)$$

$$M_\lambda = -0.15516\, X_{J,\lambda} + 0.45692\, Y_{J,\lambda} + 0.03287\, Z_{J,\lambda} \quad (3.39)$$

$$S_\lambda = 0.01608\, Z_{J,\lambda} \quad (3.40)$$

According to the requirement for the transformation:

$$Y_{J,\lambda} = L_{,\lambda+}\, M_\lambda \quad (3.41)$$

Table 3.3 shows the Smith and Pokorny fundamentals as calculated from equations 3.38–3.40.

Table 3.3 Color matching functions and chromaticity coordinates for the Smith and Pokorny fundamentals. Values are tabulated at 10 nm intervals between 400 nm and 700 nm

Wavelength	L	M	S	V_λ	l	s
400	0.0027	0.0018	0.0045	0.0045	0.6055	1.0002
410	0.0055	0.0038	0.0094	0.0093	0.5948	1.0089
420	0.0100	0.0075	0.0171	0.0175	0.5741	0.9760
430	0.0149	0.0124	0.0234	0.0273	0.5453	0.8556
440	0.0192	0.0187	0.0258	0.0379	0.5059	0.6816
450	0.0219	0.0249	0.0237	0.0468	0.4670	0.5057
460	0.0263	0.0337	0.0207	0.0600	0.4383	0.3452
470	0.0399	0.0511	0.0179	0.0910	0.4385	0.1967
480	0.0649	0.0741	0.0121	0.1390	0.4672	0.0874
490	0.1032	0.1048	0.0072	0.2080	0.4963	0.0345
500	0.1675	0.1555	0.0043	0.3230	0.5185	0.0132
510	0.2695	0.2335	0.0025	0.5030	0.5359	0.0049
520	0.3929	0.3171	0.0012	0.7100	0.5534	0.0017
530	0.4927	0.3693	0.0007	0.8620	0.5715	0.0008
540	0.5628	0.3912	0.0003	0.9540	0.5900	0.0003
550	0.6078	0.3872	0.0001	0.9950	0.6108	0.0001
560	0.6329	0.3621	0.0001	0.9950	0.6360	0.0001
570	0.6355	0.3165	0.0000	0.9520	0.6676	0.0000
580	0.6145	0.2555	0.0000	0.8700	0.7064	0.0000
590	0.5697	0.1873	0.0000	0.7570	0.7526	0.0000
600	0.5063	0.1247	0.0000	0.6310	0.8023	0.0000
610	0.4271	0.0759	0.0000	0.5030	0.8491	0.0000
620	0.3377	0.0433	0.0000	0.3810	0.8865	0.0000
630	0.2421	0.0229	0.0000	0.2650	0.9135	0.0000
640	0.1634	0.0116	0.0000	0.1750	0.9336	0.0000
650	0.1014	0.0056	0.0000	0.1070	0.9472	0.0000
660	0.0582	0.0028	0.0000	0.0610	0.9549	0.0000
670	0.0307	0.0013	0.0000	0.0320	0.9591	0.0000
680	0.0164	0.0006	0.0000	0.0170	0.9620	0.0000
690	0.0079	0.0003	0.0000	0.0082	0.9632	0.0000
700	0.0040	0.0001	0.0000	0.0041	0.9745	0.0000

The height of the LWS fundamental at its λ_{max} is 0.6373; that for the MWS fundamental at its λ_{max} is 0.3924. Figure 3.7 shows the Smith and Pokorny fundamentals, renormalized to their peak and plotted on a logarithmic axis.

A recent development is the use of the Stiles and Burch (1955, 1959) data to derive König fundamentals, e.g. (Estévez, 1979; Stockman *et al.*, 1993; Stockman and Sharpe, 2000). The goal of these attempts was to avoid the criticism that the CIE and Judd standard observers represent an amalgam of colorimetric and photometric data obtained from different individuals. A new difficulty is introduced. The Stiles and Burch pilot data do not incorporate luminosity, since the matching functions were obtained with direct energy calibrations of the primaries and test wavelengths. However, it is now considered useful for physiological fundamentals to incorporate luminosity. While the Y of the CIE 10° XYZ Standard Observer can be used to represent the relative luminous efficiency function, there are no measured dichromatic copunctal points for 10° fields (see section 3.5).

Physiologically based chromaticity diagram: A physiologically based chromaticity space was suggested as early as Maxwell (1860). He postulated an isosceles triangle with the cone fundamentals at each corner. Although inherently attractive, the Maxwell triangle is difficult to use.

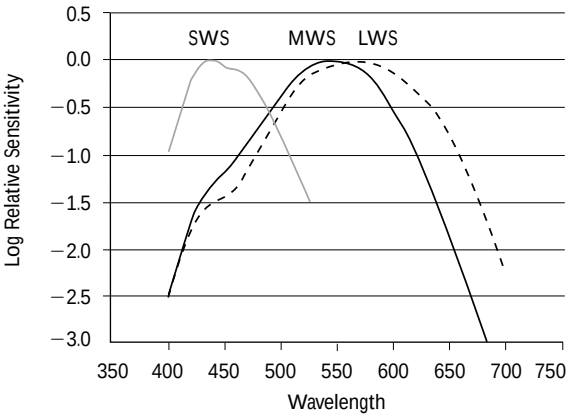


Figure 3.7 **Relative spectral sensitivity of the cone visual photopigments as proposed by Smith and Pokorny (1975). The curves are normalized to their own maxima.**

MacLeod and Boynton (1979) suggested that a useful and easily used chromaticity chart would be a constant luminance plane in which the cone spectral sensitivities formed rectangular axes. This is essentially the chromaticity diagram formed by the Smith and Pokorny fundamentals with the value of S_λ/Y_λ arbitrarily placed at unity at its peak.

$$l_\lambda = L_\lambda/(L_\lambda + M_\lambda) \quad (3.42)$$
$$m_\lambda = M_\lambda/(L_\lambda + M_\lambda) \quad (3.43)$$
$$s_\lambda = S_\lambda/(L_\lambda + M_\lambda) \quad (3.44)$$

This cone-based diagram includes the specific assumption that s_λ does not contribute to luminance. Figure 3.8 shows the cone-based diagram with s_λ plotted against l_λ. In the cone-based diagram the horizontal axis represents the exchange of LWS and MWS cone excitation at equiluminance, i.e. an increase in LWS cone excitation is offset by a decrease in MWS cone excitation but the sum is unity. The vertical axis represents variation in SWS cone excitation at a constant retinal illuminance. The protan

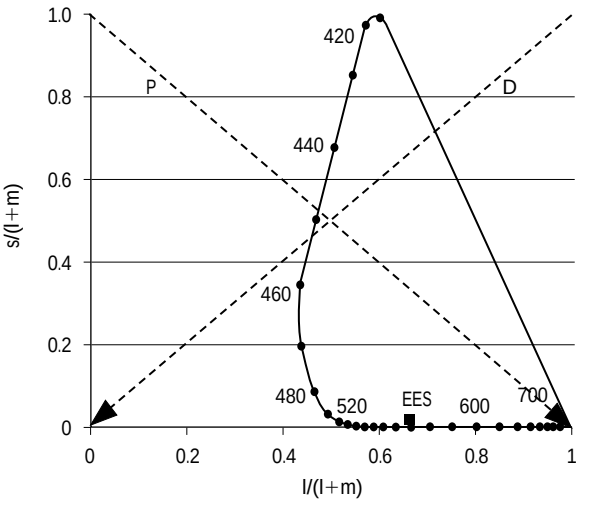


Figure 3.8 **MacLeod–Boynton (1979) diagram. Chromaticity diagram representing relative cone excitation. The relative SWS cone excitation s/(l+m) is plotted against l/(l+m), the proportion of LWS cone stimulation from equiluminant stimuli. The isochromatic lines for protanopes P and deuteranopes D converge at l=1 (point representing LWS) and m=1 (point representing MWS), respectively. Tritanopic isochromatic lines plot as parallel lines orthogonal to the ordinate.**

copunctal point is at coordinate (0,1); the deutan copunctal point is at (0,0). Tritan confusions are represented by a set of parallel, vertical lines. Although it is customary to refer to the MacLeod–Boynton diagram as a cone excitation diagram, it should be emphasized that the cone excitations are not equivalent to quantal excitation of the cones (see below).

Rod excitation relative to photopic luminance may be incorporated in the cone excitation diagram as a third dimension, by plotting $\mathbf{V}_{\mathbf{J}\lambda}/V_\lambda$ on the z axis (Pokorny and Smith, 1986). For chromaticities produced by three physical primaries, rod confusion lines for a fixed scotopic luminance will be sets of parallel lines in the equiluminance plane (Shapiro *et al.*, 1996). The angles of these lines will be determined by the particular choice of primaries.

Cone trolands: Boynton and Kambe (1980) proposed the definition of a new unit, the cone troland. Following their suggestion, we can define the L, M, and S cone trolands at wavelength, λ as:

$$L_\lambda td = (I)\, L_\lambda/\mathbf{V}_{\mathbf{J}\lambda} \quad (3.45)$$

$$M_\lambda td = (I)\, M_\lambda/\mathbf{V}_{\mathbf{J}\lambda} \quad (3.46)$$

$$S_\lambda td = (I)\, S_\lambda\, 1.6064/\mathbf{V}_{\mathbf{J}\lambda}\, (I)\, Z_\lambda/\mathbf{V}_{\mathbf{J}\lambda} \quad (3.47)$$

where I is the retinal illuminance in trolands, L_λ, M_λ, and S_λ are the Smith and Pokorny cone fundamentals, and $\mathbf{V}_{\mathbf{J}\lambda}$ is the spectral luminous efficiency of the Judd (1951) observer. The SWS fundamental is renormalized to be equivalent to the Z_λ of the Judd observer (the normalization is arbitrary in the Smith and Pokorny fundamentals and was set for convenience in the MacLeod–Boynton diagram).

The cone troland has two important properties. First, cone trolands are proportional to the quantal excitation rate of the photoreceptors. Quanta are related to trolands at wavelength λ by the formula:

$$Q_\lambda = I/\,\mathbf{V}_{\mathbf{J}\lambda}(10^7/8)(\lambda/555) \quad (3.48)$$

where Q_λ is the number of trolands, and the other symbols are described above. The quantal excitation rate for a given photoreceptor is proportional to the number of quanta multiplied by the relative absorption spectrum of the photoreceptor. For the LWS cone:

$$L_{Q,\lambda} = Q_\lambda(,L\lambda/L_{max})(\lambda_{Q,L}/\lambda) \quad (3.49)$$

L/L_{max} represents the relative spectral sensitivity of the photoreceptor and $(\lambda_{Q,L}/\lambda)$ converts from the energy base of the fundamental to the quantal base of an absorption spectrum. The term $(\lambda_{Q,L})$ is the wavelength at which the LWS absorption spectrum has its peak absorption. The peak of the absorption spectrum is shifted to shorter wavelength than the peak of the fundamental. When equations 3.48 and 3.49 are combined for each cone type:

$$L_{Q,\lambda} = L_{td,\lambda}/L_{max}(10^7/8)\,(\lambda_{Q,L}/555) \quad (3.50)$$

$$M_{Q,\lambda} = Q_\lambda M_\lambda/M_{max})\,(\lambda_{Q,M}/\lambda) = M_{ts,\lambda}/M_{max}(10^7/8)\,(\lambda_{Q,M}/555) \quad (3.51)$$

$$S_{Q,\lambda} = Q_\lambda S_\lambda/S_{max}\,(\lambda_{Q,S}/\lambda) = S_{td,\lambda}/Z_{max}(10^7/8)\,(\lambda_{Q,S}/555) \quad (3.52)$$

Thus for a given cone type, the quantal excitation at any wavelength is proportional to the cone trolands weighted by a constant.

Second, colorimetric calculations are easily made using the cone troland space (Smith and Pokorny, 1996). A light, Q specified by the Judd revised observer as x_{JQ}, y_{JQ}, Y_{JQ}, has a unique specification in the cone space given by l_Q, z_Q, Y_{JQ}. The tristimulus values, L_Q, $Y_{J,Q}$, Z_Q are calculated in the same way as for $X_{J,Q}$, $Y_{J,Q}$, and $Z_{J,Q}$. Expression in other primary systems follows the usual rules (for example, to transform between cone trolands and the phosphors of a color monitor). The equations (3.38–3.40) can be rewritten:

$$\begin{bmatrix} L_Q \\ Y_{JQ} \\ S_Q \end{bmatrix} = \begin{bmatrix} l_1/y_{J1} & l_2/y_{J2} & l_3/y_{J3} \\ 1 & 1 & 1 \\ s_1/y_{J1} & s2/y_{J2} & s_3/y_{J3} \end{bmatrix} \begin{bmatrix} Y_{JQ,1} \\ Y_{JQ,2} \\ Y_{JQ,3} \end{bmatrix} \quad (3.53)$$

where the terms l, y_{Ji}, z_{Ji} represent the specification in chromaticity coordinates for phosphor (i) and $Y_{J,\,Q,j}$ represents the luminance of light Q for phosphor (i).

3.2.6 SOURCES OF INDIVIDUAL DIFFERENCES IN COLOR MATCHING

The parameter variables that may modify color matching data were discussed in section 3.2.4. These included the field size, the luminance

level, the choice of primaries, and the method. Here we discuss the physiological mechanisms that may play a role in modifying matches for an individual or that may play a role in explaining inter-individual differences in matching. Since color matches are a linear transform of photopigment spectral sensitivity, factors that modify the spectral sensitivity can modify color matches. These factors include individual variation in photopigment spectra, variation in optical density of the photopigments, photoreceptor optics, and pre-retinal filters.

Variation in photopigment spectra: Modern molecular biology has defined the protein structure of the genes coding the photopigment opsins. To date there is at least one well-documented polymorphism of the L photopigment opsin (either the amino acid serine or alanine at position 180 on the L- opsin gene; Neitz *et al.*, 1991; Merbs and Nathans, 1992) which changes the peak wavelength of the absorption spectrum by a small amount (~3–5 nm). Sanocki, Shevell, and Winderickx (1994), using a technique which eliminates in large part inter-observer differences in prereceptoral filtering and photopigment optical density, show reliable small differences in color matching in the red–green spectral region for observers having the two different alleles. The same allele pattern also is present for the M-cone photopigment but at a low frequency and it is probable that there are other polymorphisms capable of modifying extinction spectra by small spectral shifts. There is now accumulating evidence that individuals with normal color vision exhibit small variation in the absorption spectra of the photopigments in their L and M cones.

An estimate of the extreme values that photopigment variation and optical density might assume in the normal population may be obtained by theoretically manipulating these variables and establishing the extreme values that fall within the intra-observer variability of color matching data. If WDW normalization (Wright, 1946; Wyszecki and Stiles, 1982) is used for both data and the synthesized CMFs, the effects of individual differences in prereceptoral filtering are eliminated. The choices of primaries and normalizing wavelengths are unimportant as long as the same are used for both the data and the synthesized CMFs. Smith, Pokorny, and Starr (1976) performed such an analysis for the Stiles and Burch (1955) 2° color matching data and a set of theoretical spectra based in large part on a transformation of the Judd (1951) color matching functions. The analysis can define the potential extreme limits of normal variation but cannot specify which variable or combination of the variables is responsible for the normal variation. Calculated variation in the S, M, and L cone spectral positions indicated the data variability was dominated by the theoretical L and M cone spectral shifts; S cone shifts had relatively little effect. Wavelength shifts in a range from −4 to +2 nm in the M cone spectrum and −3 to +7 nm in the L cone spectrum predicted WDW unit coordinates that fell within the extremes of the Stiles data.

Effective optical density of the photopigments: Light must be absorbed in the photopigments to be seen. The concentration of the pigment and the length of the light path affect an absorption spectrum. As light traverses a greater concentration of pigment, more light is absorbed and the extinction spectrum broadens. The rule governing absorption follows Beer's law. In the case of a visual photopigment, expressed in decadic base:

$$A = (1–10^{-\varepsilon\lambda^{cl}}) \qquad (3.54)$$

where A is the fraction absorbed, ε_λ is the decadic extinction coefficient, c refers to the concentration and l to the pathlength. The extinction coefficient, ε_λ, is a wavelength-dependent function that is characteristic of the photopigment. For visual photopigments, concentration and pathlength can be grouped as a single factor, the effective optical density of the photopigment.

A graphical sketch to demonstrate how the absorption spectrum broadens as the effective optical density increases is shown in Figure 3.9. The insert in the top panel schematically illustrates a path through a beaker of the rod photopigment rhodopsin. The curves show the calculated effective fractional absorbance at each successive layer of pigment. Light reaching each successive layer is the product of the incoming

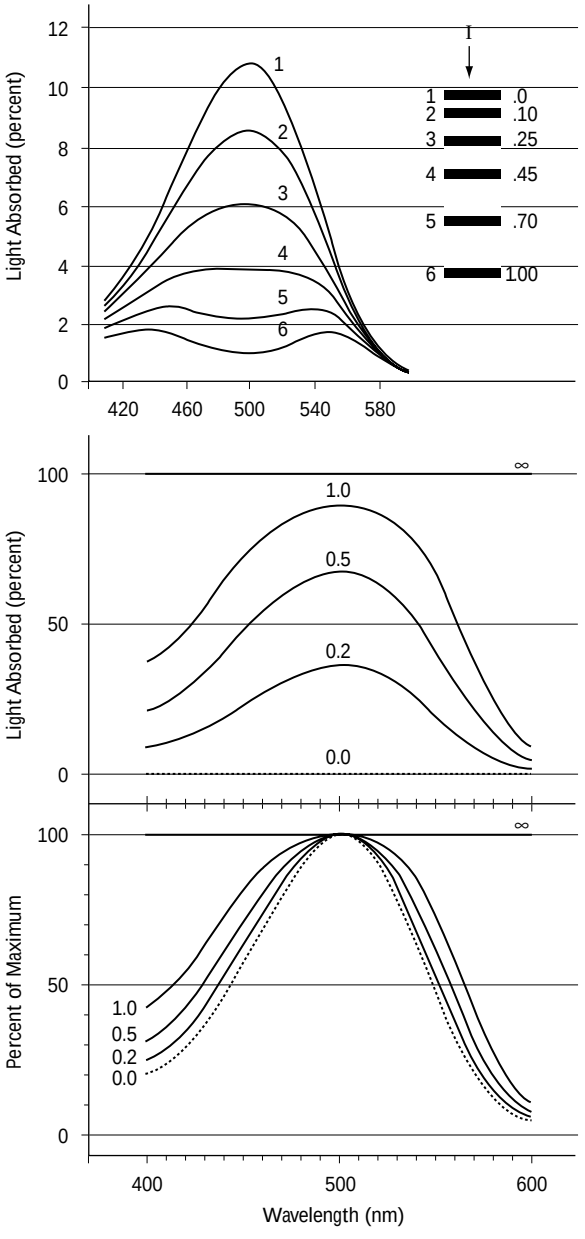


Figure 3.9 **Effect of optical density on spectral sensitivity. (The top panel is from Goldstein and Williams, 1966; the bottom two panels are from Dartnall, 1957.)**

light and the transmission of the layers that precede them. At the initial layer (1), the absorption resembles that of the extinction spectrum. At the final layer (6), most of the absorption is taking place in the tails of the extinction spectrum. The absorptance spectrum is the sum of the absorbances of all layers. The middle panel shows the fractional absorption spectrum, calculated for different optical densities of pigment. The lower panel shows the relative fractional absorption.

The analysis proposed by Smith, Pokorny and Starr can also be applied to variation in optical density. Calculated variation in photopigment optical densities could also predict the variability of the Stiles and Burch data. In this analysis a higher optical density was required for the L cone photopigment, a result found for all published comparisons of L and M cone optical density (Miller, 1972; Smith and Pokorny, 1973; Bowmaker *et al.*, 1978; Wyszecki and Stiles, 1980; Burns and Elsner, 1993). An optical density range of 0.15 to 0.45 for the M cone spectrum and 0.25 to 0.55 for the L cone spectrum predicted WDW unit coordinates that fell within the extremes of the Stiles (1955) data. The estimated standard deviation was 0.05.

Cone photoreceptors vary in length, with retinal position being longest in the center of the fovea (Polyak, 1957). Field size is an important variable in matching which may be attributed partially to changes in the effective optical density of photopigments. Pokorny, Smith and Starr (1976) extended their analysis to the difference of the Stiles (1955) 2° and the Stiles and Burch (1959) 10° data. The differences between the mean 2° and 10° data could be explained by a reduction in optical density of 0.10–0.15. Pokorny, Smith and Starr also calculated the theoretically acceptable range of optical densities of the L and M cone photopigments for data collected with a small field at several eccentricities. Thomson and Wright (1947) obtained small field tritanopic data from W.D. Wright's eye with a 15′ field fixated directly or placed at eccentricities of 20′ and 40′ from fixation. There have been a number of experimental verifications of the change in color matches with field size (Horner and Purslow, 1947; Pokorny and Smith, 1976; Burns and Elsner, 1985; Eisner *et al.*, 1987; Swanson and Fish, 1996). These studies are all consistent with the interpretation that the effective optical densities of the photopigments decrease as field sizes increase. Figure 3.10 shows the estimated L cone optical density as a function of field size from a number of studies.

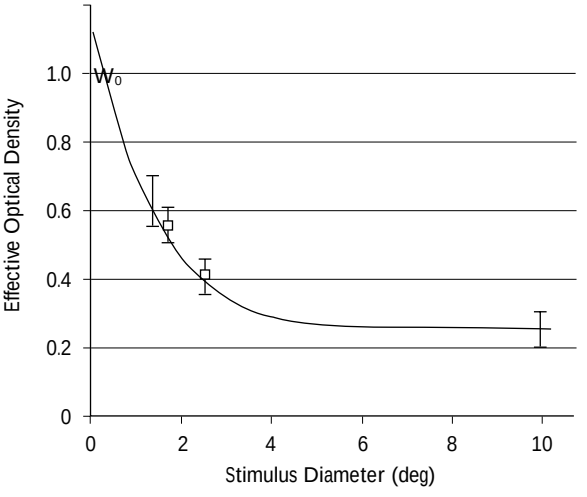


Figure 3.10 **Optical density of the L cones as a function of field size. (From Pokorny *et al.*, 1976.)**

Pre-retinal filters: The other major sources of variation in color matching are pre-retinal filters, namely the lens and macular pigment density. Figure 3.11 shows the density spectra for an average lens of a 32-year-old observer and the average macular pigment for a 2° field. These both show substantial variation in the population.

There is significant inter-individual variation in lens transmission and also a substantial change with age. Van Norren and Vos (1974) calculated from Crawford's (1949) scotopic spectral sensitivity data that individual variation in ocular absorption for a group of 50 young observers (17–30 year) is about ±25% of the average absorption at short wavelengths. Adult lens transmission has been characterized as having two components, T_{L1} and T_{L2} with T_{L1} varying with age (Tan, 1971; Pokorny *et al.*, 1987). An algorithm that allows calculation of an average lens transmission between the ages of 20 and 60 years is:

$$T_L = T_{L1}\,[1 + 0.02(A - 32)] + T_{L2} \quad (3.55)$$

where T_L, T_{L1}, and T_{L2} represent the spectral lens densities tabulated in Table 3.4 and A is the chronological age. Beyond the age of 60, there is an acceleration in decrease in lens (Pokorny *et al.*, 1987). This acceleration is characteristic of a population average but was not observed in studies where observers were screened so as to eliminate those with incipient cataracts (Weale, 1988; Moreland *et al.*, 1991).

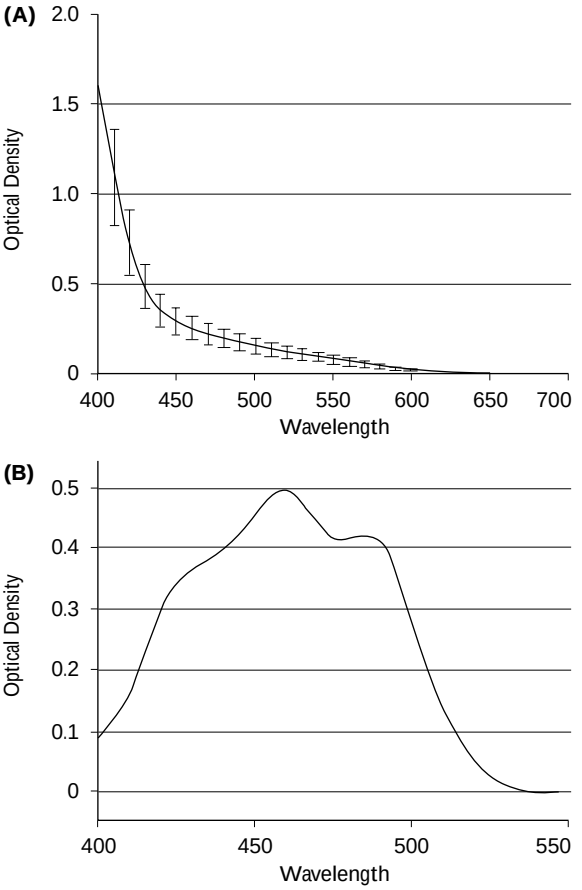


Figure 3.11 **(A) The optical density of the lens (van Norren and Vos, 1974) and (B) the macular pigment (Wyszecki and Stiles, 1982).**

The macular region of the human retina contains a non-photolabile pigment that selectively filters light arriving at the receptors. Entopic viewing of Maxwell's spot (section 3.2.4) is interpreted as a visualization of the spatial distribution of macular pigment (Palmer, 1978). For many observers Maxwell's spot is seen to be darkest in the center of the fovea but for others there is considerable spatial structure (Miles, 1954).

Studies of the optical density of the macular pigment have used widely different methodologies. Since macular pigment varies with field position (see below), it is necessary when evaluating literature reports to take into account the field size employed in the various studies. Perhaps the most striking feature of the literature reports

Table 3.4 Tabulation of the Optical Density of the Total Lens Transmission Function T_L for an Average 32-year-old Observer and Separation of T_L into Components: T_{L1} Represents Portion Affected by Aging after age 20, and T_{L2} Represents Portion Stable after age 20.* (After Pokorny, Smith and Lutze, 1987)

Wavelength	Optical Density			Wavelength	Optical Density		
nm	T_L	T_{L1}	T_{L2}	nm	T_L	T_{L1}	T_{L2}
400	1.933	0.600	1.333	530	0.120	0.120	–
410	1.280	0.510	0.770	540	0.107	0.107	–
420	0.787	0.433	0.354	550	0.093	0.093	–
430	0.493	0.377	0.116	560	0.080	0.080	–
440	0.360	0.327	0.033	570	0.067	0.067	–
450	0.300	0.295	0.005	580	0.053	0.053	–
460	0.267	0.267	–	590	0.040	0.040	–
470	0.233	0.233	–	600	0.033	0.033	–
480	0.207	0.207	–	610	0.027	0.027	–
490	0.187	0.187	–	620	0.020	0.020	–
500	0.167	0.167	–	630	0.013	0.013	–
510	0.147	0.147	–	640	0.007	0.007	–
520	0.133	0.133	–	650	0.000	0.000	–

* The optical density of the lens of an average observer between the ages of 20 and 60 years old may be estimated by

$$T_L = T_{L1}[1 + 0.02(A\text{-}32)] + T_{L2},$$

For an average observer over the age of 60

$$T_L = T_{L1}[1.56 + 0.0667(A\text{-}60)] + T_{L2},$$

where A is the observer's age.

T_L is the van Norren and Vos (1974) tabulation of lens density scaled by 1.333 to represent a 32-year-old observer (the average age of the Stiles and Burch observers) with a small pupil (<3mm). To estimate the lens density function for a completely open pupil (>7mm), multiply the tabulated values by 0.86.

For the Wyszecki and Stiles (1982) lens density function, the following values may be substituted:

Wavelength	Optical Density		
nm	T_L	T_{L1}	T_{L2}
400	1.600	0.600	1.000
410	1.093	0.510	0.583
420	0.733	0.433	0.300

is the marked individual variation in macular pigment; for a 2° field size, individuals may vary from 0.0 to >1.0 optical density at 460 nm, the wavelength of peak absorption (Vos, 1972; Pease *et al.*, 1987; Werner *et al.*, 1987).

Psychophysically measured density of the macular pigment decreases with increase in field size. Estimates from literature reviews of the average macular pigment optical density for a 2° field vary from 0.35 (Vos, 1972) to 0.50 (Wyszecki and Stiles, 1982). Stiles (1955) reported that the 10° colorimetric data are characterized by macular pigment densities of about 0.25 times the 2° optical density. Four studies have attempted to characterize the retinal distribution of macular pigment by comparison of data taken for a small field at a number of retinal eccentricities (Ruddock, 1963; Stabell and Stabell, 1980; Viénot, 1983; Hammond *et al.*, 1997). These data may be characterized with an exponential decrease in macular pigment optical density with retinal eccentricity (Moreland and Bhatt, 1984; Hammond *et al.*, 1997).

3.3 CHROMATIC DETECTION

Detection refers to the ability to recognize a change in light level. The threshold for detection can be measured in the fully dark-adapted eye (absolute threshold) or on a background (increment threshold). The threshold represents some criterion change presumably correlated with a change in a neural response from its steady-state adapting level. It is possible to specify absolute threshold in terms of the number of quanta

needed for detection. The increment threshold is usually expressed either as the increment or change in quanta, ΔQ, from the background level or as the contrast ratio of the increment to the background, ΔQ/Q. Increment thresholds may equally well be expressed in quantal, radiance or luminance units.

Sensitivity is inversely related to detection; it is the reciprocal of the quanta at the detection threshold. The human visual system maintains an approximately even state of contrast sensitivity over a billion-fold range of light levels. Detection thresholds can be measured over this entire range of illumination. In the mid-periphery of the dark-adapted eye, where the visual system is most sensitive, the absolute detection threshold requires only a few quanta. The visual system continues to maintain sensitivity even at steady-state background levels of $10^{11.5}$ quanta (about 20 000 td at 555 nm). At this quantal excitation level, there is a 50% depletion of the MWS and LWS cone photopigments. At such levels, an increment of many millions of quanta is needed to detect a change from the steady excitation level.

Part of the sensitivity range of the human visual system is obtained by the duplex nature of the retina. There are two sub-systems each with an operating range over a million-fold, that show overlapping ranges of sensitivity. One receptor system, the rods, constitutes a high sensitivity, low visual acuity, color-blind system. The other system, the cones, represents a high threshold, high visual acuity, color system. There are a variety of sources which treat the topics of detection and sensitivity more fully (Cornsweet, 1970; Barlow, 1972; Hood and Finkelstein, 1986). Here we briefly review some explanatory concepts in detection and review chromatic detection data aimed at isolating cone mechanisms and/or postreceptoral channels.

3.3.1 THRESHOLD-VERSUS-RADIANCE (TVR) FUNCTIONS

Detection thresholds can be measured as a small, brief increment, ΔR on a larger steady background of radiance R. At each background level, the observer first adapts to the background. If the test and background are foveally fixated and the spectral distribution is uniform, detection will be mediated by cone mechanisms. The data may be expressed in several different metrics. The detection threshold can be plotted as the log ΔR versus log R and the function is called a threshold-versus-radiance (TVR) function. Alternately, the data may be expressed in units of retinal illuminance and the function is called a threshold-versus-illuminance (TVI) function. Data may also be specified in quanta/deg^2/sec. The TVR function shows a characteristic shape. At low radiances, detection threshold is unchanged from the absolute threshold on a zero background. As the background is increased, threshold starts to increase and reaches a limiting ΔR/R slope of unity. At this limiting slope, sensitivity is said to be in the Weber region; threshold is a constant percentage of the background, or a constant contrast.

3.3.2 EXPLANATORY CONCEPTS IN DETECTION

In the psychophysical literature, there have been three important explanatory concepts in describing the TVR function. These are adaptation, saturation, and noise.

3.3.2.1 Adaptation

Adaptation refers to mechanisms that expand the dynamic range of response of the system as a whole. Any given retinal element has a limited response range of perhaps 400-fold. Thus, without adaptation, as light level increases the range available for greater stimulation decreases. Following adaptation, a neural element may maintain this 400-fold response range for a wide range of background light levels. The literature distinguishes two major sub-types of adaptation. Subtractive adaptation resets the steady-state signal without affecting the response to a stimulus pulse. Multiplicative adaptation scales both the responses to the steady-state and to a stimulus pulse. Multiplicative adaptational mechanisms are sometimes called gain controls.

Multiplicative adaptation has many sources. The reflex response of the closing of the pupil in bright illumination is a source of multiplicative adaptation in the natural environment. Pupil constriction reduces the amount of retinal illumination from both a steady-state background and from a light increment. Pupil constriction

allows a 16-fold increase in the total operating range of the visual system, under conditions of natural viewing and thus plays a relatively minor role in adaptation in the natural environment. In threshold experiments, the effective pupil is often held constant by use of an artificial pupil (Troland, 1915; Pokorny and Smith, 1997) allowing the study of neural adaptational mechanisms. An artificial pupil is an aperture smaller that the natural pupil which the observer looks through. An alternate technique sometimes used in Maxwellian view systems places a limiting aperture in a plane conjugate to the natural pupil.

Photopigment depletion is another source of multiplicative adaptation. Breakdown and regeneration of photopigment are reciprocal processes. At low illuminations, regeneration works to maintain a full complement of photopigment. At high illuminations, the visual photopigment is depleted (sometimes termed bleaching). In steady illumination, there is no net effect. For a first order kinetic process (Rushton, 1972) the amount of photopigment depleted (1−p) is given by:

$$(1-p) = I/(I + I_o) \qquad (3.56)$$

where p is the amount of photopigment available, I is the retinal illuminance and I_o is a constant, the illumination at which 50% of the photopigment is depleted. For cone photopigments at their λ_{max}, the value of I_o is $10^{11.5}$ quanta (about 20,000 td at 555 nm). At levels above 20,000 td, there is sufficient depletion that only a lower percent of the incident background quanta are absorbed. If the system is responsive at low bleaching levels, it will remain responsive at higher levels since further increases in incident background quanta will be offset by the reduced probability of absorption. This mechanism of adaptation usually plays only a small role in the natural environment, however, snow fields and sunlit beaches may provide sufficient illumination to allow substantial bleaching.

The most important sources of multiplicative adaptation occur neurally. Recordings from primate horizontal cells, second-order neurons in the retina show adaptation over a wide range of light levels (Smith *et al.*, 2001) though the extent of adaptation is not as great as found later in the retina (Lee *et al.*, 1990), or psychophysically.

It is common to lump sources of multiplicative adaptation in a single term. At a computational level, multiplicative adaptation is a scalar less than one that modifies the adapted response, R_A to both the background, Q_A and the increment, ΔQ. The response to the background is given by:

$$R_A = Q_A/(1 + kQ_A) \qquad (3.57)$$

where k is an adaptation constant. Provided that the threshold increment does not perturb the adaptation level, the increment threshold is described by the equation:

$$\Delta Q = (Q_{th})\,(1 + kQ_A) \qquad (3.58)$$

where (Q_{th}) is the absolute threshold, determined by criterion, noise, and/or quantal requirements. The adaptation constant, k, is the reciprocal of the quanta at which threshold is raised two-fold. On a double logarithmic scale, the limiting slope is unity, as demanded in the Weber region. This equation describes the data of TVR functions obtained for the cone mechanisms on large steady backgrounds. The function is sketched in Figure 3.12.

Subtractive adaptation is considered to remove part of the adapting background (Geisler, 1981; Adelson, 1982; Hayhoe *et al.*, 1987). When subtractive adaptation is added to

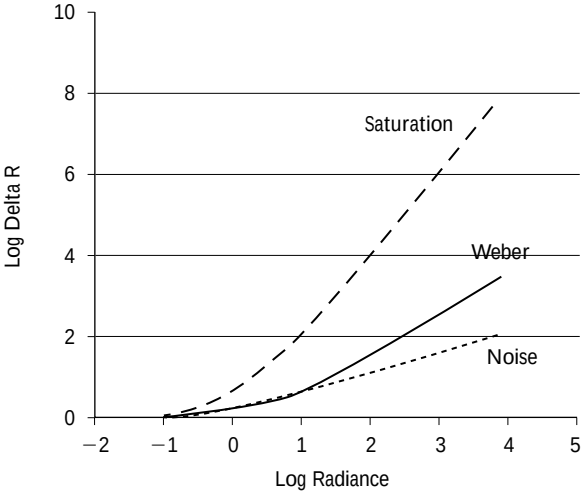


Figure 3.12 **The increment threshold functions predicted by three types of detection mechanism, a saturation function, adaptation following Weber's law, and a noise mechanism.**

a process already controlled by multiplicative adaptation, the effect is to shift the TVR function on the horizontal axis.

3.3.2.2 Saturation

Any given neural element has a limited response range of perhaps 400-fold. The response of a non-adapting retinal element initially increases with the incident quanta, but with diminishing slope. At a high level above threshold, the response is constant and is said to have saturated. In vision, the saturating response, R is usually described by a Naka–Rushton equation:

$$R = R_{max} (Q)/(Q + Q_{sat}) \quad (3.59)$$

where R_{max} is the maximal response, Q is the incident quanta and Q_{sat} is the semi-saturation, or number of quanta at which the response is half the maximum value. It is clear from equation (3.60) that sensitivity is greatest at low light levels, and that at high light levels where the response has saturated, a detectable increment in response will not be measurable. The equation for detection for a saturating system obeying the Naka–Rushton equation is:

$$\Delta R = (\delta/R_{max}) (Q + Q_{sat})^2/(Q_{sat} - (\delta/R_{max}) (Q + Q_{sat})) \quad (3.60)$$

where δ is the criterion, R_{max} is the maximal response, Q is the incident quanta and Q_{sat} is the semi-saturation, or number of quanta at which the response is half the maximum value. Normally the term (δ/R_{max}) is very small. At absolute threshold, where Q is zero:

$$\Delta R \sim (\delta/R_{max}) (Q_{sat}) \quad (3.61)$$

At higher quantal excitation levels where $Q >> Q_{sat}$:

$$\Delta R \sim (\delta/R_{max}) (Q^2) \quad (3.62)$$

This equation then describes a function that has a slope of two, steeper than the limiting slope of the Weber function. For some values of δ, R_{max} and Q_{sat}, the denominator of equation 3.60 becomes negative, implying that negative light is required at threshold. It is at this point that the system is said to have saturated. Equation 3.60 is sketched in Figure 3.12.

Saturation can be found in cone-mediated TVR functions under special conditions of measurement. Protocols that reveal saturation include pulsed backgrounds (King-Smith and Webb, 1974; Shevell, 1977; Stockman *et al.*, 1993) and the probe-flash technique (Hayhoe *et al.*, 1987; Hood, 1998). A large steady background controls the overall state of adaptation. An adapting pulse of a second or so, called the flash is superimposed on the background, and a brief test pulse, the probe is superimposed on the adapting flash. The probe may be presented at any time interval during the flash presentation. If the probe is presented in the middle of the flash duration, threshold follows a Weber function, indicative of neural adaptation to the flash. If the probe is presented at flash onset, threshold is elevated above the Weber function demonstrating a slope of two before saturation, indicative of a saturating process. The probe-flash technique has been used to delineate a different time course for subtractive and multiplicative adaptation (Geisler, 1981; Hayhoe *et al.*, 1987).

Mechanisms of neural adaptation: The neural multiplication and neural subtractive adaptation inferred from the data described above has computational importance. Another approach is to ask what kind of neural mechanism might exist. Neural mechanisms may be feed-forward or feed-backward. Historically, feed-forward is considered unsatisfying since it must be subject to saturation.

In **subtractive feedback**, a portion, k of the signal is subtracted from the signal. In the steady state subtractive feedback is described as

$$R = Q - (kR) \quad (3.63)$$
$$R = Q/(1+k) \quad (3.64)$$

Subtractive feedback acts to shift the start of a saturating mechanism to a higher radiance level. Models of subtractive feedback may postulate that the feedback is developed slowly in time compared with the signal. Slow subtractive feedback shows pronounced overshoots following a rapid increase in radiance level.

In **divisive feedback** the signal is divided by a portion of the signal. In the steady state response is given by:

$$R = Q/(1+kR) \quad (3.65)$$
$$R = -\tfrac{1}{2}k + [(1+4kQ)/4k^2]^{0.5} \quad (3.66)$$

In divisive feedback, the response increases as the square root of the signal; i.e. only part of the signal is removed by the adapting mechanism. Subtractive feedback acts to shift a saturating mechanism by the square root of the steady-state radiance.

Combined neural mechanisms: The subtractive and divisive feedbacks described above do not produce Weber behavior. There is interest in the literature as to how Weber behavior may be produced. One possibility, as proposed by Koenderink, van de Grind and Bouman (1970) was to weight a subtractive feedback pathway with a feed-forward signal that represented the average radiance over time.

3.3.2.3 Noise

An important concept in detection is that the signal may be noisy. There are two consequences of the noise concept in visual thresholds. First, noise affects the slope of the psychometric function and, second, a noise-limited process affects the slope of the TVR function. The applicability of the noise concept at absolute threshold followed the recognition that the statistics of light emission in a brief pulse follow a Poisson process. The introduction of the noise concept in visual processing occurred at a time when engineers were developing methods to assess signal transmission in electronic transmission, i.e. signal detection theory. A detection event then requires that the response to a signal on a noisy background be some criterion amount larger than the noise level alone. If the noise is Poisson-limited, its variance is given by the mean. Suppose threshold is defined at 55% correct detection, i.e., detection occurs at one standard deviation above the noise. The increment threshold is defined as:

$$\Delta R = (Q^{0.5}) \quad (3.67)$$

In the classic paper of Hecht, Shlaer, and Pirenne (1942), the slope of the psychometric function at rod absolute threshold was related to the uncertainty in the number of quanta delivered per light pulse. An important and undisputed conclusion from this study was that a single quantum is sufficient to excite a photoreceptor. However, the study also suggested that on average five to eight absorptions were required for detection on 55% of the trials. Thus detection is limited not by absorption but by some neural criterion. Barlow (1956) introduced the idea that the photoreceptor responses might themselves be noisy in the absence of stimulation. This idea is sometimes called 'equivalent noise,' EN. Equation (3.60) can be rewritten:

$$\Delta R = (EN + Q)^{0.5} \quad (3.68)$$

In the absence of a real background, absolute threshold will be limited by the equivalent noise. Once the background noise exceeds the neural noise, detection will be limited by the background noise. The limiting slope of a noise-limited TVR function is 0.5 (see Figure 3.12).

The noise concept has been most important at absolute threshold. As the steady-state light level increases, multiplicative adaptation processes take over and detection is in the Weber region. Provided the multiplicative adaptation process follows the source of the noise, it will reduce the effect of both the average background and its variance. Thus, the level of stimulus noise or early neural noise will approach a constant value once gain mechanisms are active. It is also recognized that threshold may be limited by late neural noise that follows gain mechanisms. There are data that suggest that the intrinsic noise of spike generation at the retinal ganglion cell is constant at different superthreshold illuminations (Troy and Robson, 1992; Troy and Lee, 1994).

There are conditions for which cone-mediated detection obeys equation (3.68) (Barlow, 1957; Bouman and Koenderink, 1972). The conditions to obtain noise-limited detection in cones involve the use of very small test and adaptation fields. These conditions probably restrict the role of neural adaptation.

3.3.3 DETECTION ON SPECTRAL BACKGROUNDS

The TVR function for foveally fixated achromatic targets measured with large steady backgrounds and small test stimuli follows the form of

equation (3.58). However, when chromatic backgrounds and tests are used, multiple branches or partial components of multiple TVR functions can be measured depending on the choice of test and background wavelengths. Further, the spectral sensitivity of the detection mechanisms varies, depending on the background wavelength. These results were attributed to multiple underlying cone mechanisms (i.e. a photoreceptor and its adaptation mechanism). An early goal of these studies was to separate and measure the spectral sensitivity of the cone mechanisms with the view that they would reveal the underlying visual photopigments.

One important pre-theoretical concept was that a spectral background might exert a differential effect on the three classes of cones. The differential spectral sensitivity of the cone photopigments might be exploited to allow isolation of each cone mechanism by an appropriate choice of background wavelength e.g. Wald (1964). This idea in turn originated in the photochemical studies of extracted pigments (Dartnall, 1957). It was recognized that the constituents of extracted photopigment mixtures could be studied by partial bleaching, i.e. by bleaching with a spectral wavelength that exploited the differential transmissivity of the pigment mixture. In human psychophysics, it became clear that this exploitation could be achieved at levels below photopigment depletion. These data suggested that the neural adaptation mechanisms of the cone mechanisms were equivalently independent. The major development of the relation between cone-mediated TVR functions and the underlying cone mechanisms was performed by Stiles (1959, 1978).

3.3.3.1 Displacement laws

When TVR functions are measured using monochromatic lights for test and background stimulus, there are two important laws which govern the placement of the TVR function on the horizontal or vertical axis. Consider a single mechanism consisting of a photopigment with neural elements that show multiplicative adaptation, giving a classical TVR function. From equation (3.58), we see that two independent factors govern the TVR function. The term Qth for absolute threshold determines the vertical intercept and the term k for adaptation determines the transition to the Weber region.

The vertical displacement law (test wavelength variation): Suppose test and background are chosen near the peak absorption of the photopigment. At absolute threshold, the mechanism is at its most sensitive, since the chosen wavelength allows the highest probability of absorption. If the test wavelength is now changed, the absolute threshold will be higher, since absolute threshold is determined partially by the reciprocal absorptivity of the underlying photopigment (Chapter 2). The transition to the Weber region, however, is determined by the relative absorptivity of the photopigment at the adapting background that remains unchanged. The TVR functions for different test wavelengths will be displaced vertically, but will make their transition into the Weber region at the same background radiance. The vertical displacement law states that variation in test wavelength can only affect the vertical position of the TVR function for a single mechanism.

The horizontal displacement law (background wavelength variation): In comparison, consider variation in the background wavelength, leaving the test wavelength fixed near the peak absorption of the photopigment. The value of absolute threshold is not affected by the choice of background wavelength. However, the effectiveness of the background, Q_A is determined by the absorption probability of the photopigment. The transition to the Weber region occurs at higher radiances for background wavelengths other than at the peak absorptivity of the photopigment. The TVR functions will be displaced horizontally on the radiance axis but will share the same vertical position since they are pinned at absolute threshold. The horizontal displacement law states that variation in background wavelength can only affect the horizontal position of the TVR function for a single mechanism.

It is clear that the spectral sensitivity of a single mechanism can be assessed in two different ways. Spectral sensitivity can be assessed by using a background of fixed wavelength and a test stimulus of varying wavelength. The background radiance is fixed and test radiance is

varied to achieve increment threshold. This is called a test sensitivity function. Alternatively, spectral sensitivity can be assessed by using a test stimulus of fixed wavelength and a background stimulus of varying wavelength. The radiance of the background is varied to achieve a 10-fold rise in test wavelength threshold above absolute threshold. This spectral sensitivity function is called a field sensitivity function. The two methods each have advantages and disadvantages. The test sensitivity method is more straightforward, and requires minimal data collection (one threshold per test wavelength). Test sensitivity has seen the greater use (Wald, 1964; Eisner and MacLeod, 1981; Yeh *et al.*, 1989). Its disadvantage is that the human cone mechanisms are highly overlapping, while the success of the technique depends on a large differential sensitivity between constituent mechanisms. It is virtually impossible to choose a background wavelength that will yield test spectral sensitivity for the entire spectrum for a given cone mechanism. The field sensitivity method is more data-intensive; a TVR function must be measured at each background wavelength. The overlap of the human cone spectral sensitivities may result in overlap of the TVR functions. Thus, a value for the spectral sensitivity at 10× absolute threshold may in some cases depend on extrapolation of a fitted TVR function rather than a direct measurement.

The π mechanisms: If the cones have independent adaptation pathways, then the displacement laws will hold for each independent mechanism. Additionally, provided the adaptation mechanism is such as to generate a TVR function showing Weber's law, then the spectral sensitivity of the underlying mechanism may be inferred even if only a small segment of the TVR function is measurable. It is also possible to test the independence of the adapting mechanisms, by superimposing backgrounds that are mixtures of component wavelengths. Stiles developed the rationale of the displacement laws and the field sensitivity technique. Stiles used a consistent experimental paradigm. The background was 10° and the test was a 1° by 1° square pulsed for 100–200 msec. Stiles obtained a number of functions of differing spectral sensitivity, associated with different combinations of test and field wavelength. He termed these π mechanisms (Figure 3.13). These were subsequently replicated and subjected to tests of independence (Pugh and Kirk, 1986).

***π* 1,2,3:** The first three Stiles π mechanisms are associated with the SWS cone mechanism. The π 2 mechanism is the most sensitive mechanism, obtained at absolute threshold. The π 2 mechanism occurs variably among observers, and thus has not been the subject of extensive analysis. The π 1 mechanism is obtained at low to moderate adapting levels. The π 3 mechanism is obtained when adaptation fields are close to bleaching levels for the LWS and MWS cones. One difficulty in obtaining the π 1 and the π 3 mechanisms is that even with a 435 nm test flash near the peak of the SWS cone spectral sensitivity, an auxiliary long wavelength conditioning field is needed to suppress the other cone types. The finding of multiple mechanisms for SWS cone function was disappointing in that it revealed a failure to isolate a unitary mechanism that could be associated with a presumed SWS cone photopigment. Further, the long wavelength sensitivity slopes of the π 1, the π 2, and the π 3 mechanisms were not consistent with typical absorption functions for photopigments. These findings led Stiles to realize that his techniques did not reveal a unitary adapting

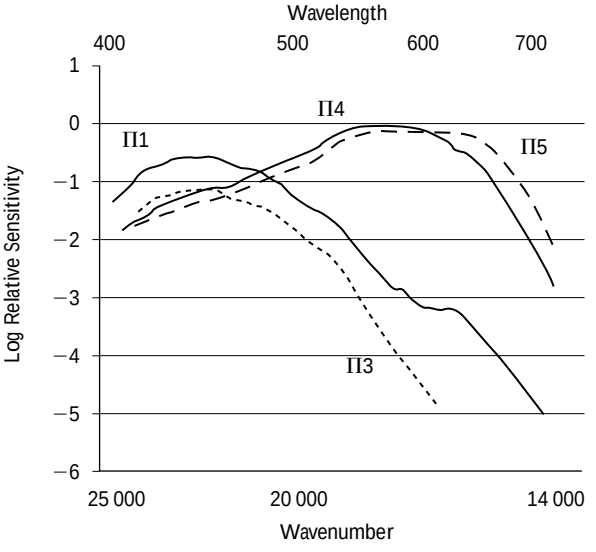


Figure 3.13 **Stiles field mechanisms. Log relative quantal field sensitivities of the π mechanisms plotted as a function of wavenumber.**

mechanism. Subsequent studies provided further indication that the π 1 to π 3 are not consistent with a unitary adapting mechanisms (Pugh, 1976; Mollon and Polden, 1977; Pugh and Mollon, 1979). In particular, there were failures of additivity when the adapting field was composed of short wavelength and long wavelength components. When the background was composed of low luminance short wavelength and long wavelength components, threshold was raised more than predicted from the additivity of the component fields (Pugh, 1976). When the background was composed of a high luminance short wavelength and lower luminance long wavelength components, threshold was raised less than predicted from the additivity of the component fields. Another irregularity of the 1 to π 3 mechanisms was called the 'limited conditioning effect.' As a long wavelength background was raised, threshold for a 435 nm test flash rose for about 0.7–0.8 log unit, and then stabilized, i.e. showed no further effect of an increase in the adapting background. The threshold started to increase again only when the background reached a level capable of bleaching the LWS and MWS cones. The final abnormality of the π 1 mechanism was one of temporal adaptation and termed 'transient tritanopia.' Following extinction of a long wavelength background, threshold for a short wavelength test flash first rose precipitously and then recovered very slowly. In comparison following extinction of a short wavelength background that caused the same steady threshold increment, sensitivity fell rapidly. A model of the π 1 and π 3 pathways was suggested by Pugh and Mollon (1979) in which the SWS cone was subject to two independent sources of adaptation. Adaptation could occur both in the SWS cone and at a later 'second' site where there was spectral opponency between SWS cones and summed activity of LWS and MWS cones (see Chapter 6 on spectral opponency in the retina). The failure of the independence tests for π 1 was attributed to independence of first site and second site adaptation. The limited conditioning effect of long wavelength adapting fields was attributed to stabilization of the steady-state response at the second site at bleaching levels of the LWS and MWS cones. Transient tritanopia was attributed to the adaptational properties of the second site. In modern models of adaptational mechanisms, the gain controls are carried forward to subsequent sites. An early stage of multiplicative adaptation serves to stabilize the later opponent sites and thus the first and second sites cannot be truly independent. The adaptational abnormalities of the π 1 mechanism are not predicted if the SWS cone shows Weber adaptation preceding the second site. The Pugh and Mollon model can be rephrased if we consider that the SWS cone pathway is subject only to partial adaptation at the first site. There is adaptation at the second site that acts to subtract most of the opponent signal.

π 4 and π 5: The remaining Stiles π mechanisms are associated with detection mediated primarily by MWS and LWS cones. These mechanisms obey the displacement laws and fulfil many criteria for unitary mechanisms. The mechanism Stiles termed π 4 has often been associated with the MWS photopigment and to a lesser extent π 5 has been associated with the LWS photopigment. Nonetheless, other evidence suggested that π 4 and π 5 did not represent photopigment sensitivities. This evidence came from many sources: The π 4 and π 5 were broader than the spectral sensitivities of X-chromosome linked dichromats believed to have only one of these mechanisms (Boynton, 1963). The π 4 and π 5 mechanisms showed unexpected interactions that were not consistent with unitary mechanisms (Boynton, Ikeda and Stiles, 1964). If the parameters of the test size and duration are changed, the spectral sensitivities of the isolated mechanisms do not necessarily agree with those of the π 4 and π 5 mechanisms (Ingling and Martinez, 1981; Stockman and Mollon, 1986). Finally, test sensitivity measurements yield narrower functions than the classical Stiles field mechanisms (Wald, 1964; Eisner and MacLeod, 1981; Yeh *et al.*, 1989; Stockman *et al.*, 1993).

Today it is conceded that the π mechanisms do not represent isolated cone spectral sensitivities. With the acceptance of the parallel pathway description of the retina (Chapter 6), it is recognized that the choice of spatio-temporal parameters will favor detection either in PC- or in MC-pathways. Thus, obedience of the displacement laws may represent isolation of a higher order neural mechanism by choice of the

spatio-temporal presentation, rather than isolation of a single photoreceptor by adaptation. The spatio-temporal parameters of the stimulus will determine the relative sensitivities of the MC- and PC-pathway channels. If these sensitivities are similar, probability summation among pathways should be considered. The choice of a 100–200 msec square-wave pulse is one that favors both MC- and PC-pathway channels. In comparison, a brief 5 msec pulse would favor the MC-pathway while a 1 second gaussian-shaped pulse would favor the PC-pathway.

3.3.4 INCREMENT DETECTION ON WHITE

With a white-adapting field, a test sensitivity function may be obtained as a function of test wavelength for a large, long test pulse. This function shows three peaks, separated by two notches (Sperling and Harwerth, 1971). There is a prominent peak near 450 nm with a notch near 500 nm. There are peaks at 540 and 600 nm, separated by a notch near 570 nm. The prominence of the peaks depends on the spatio-temporal characteristics of the test stimulus and the size and luminance of the background (Figure 3.14). The most prominent peaks are obtained when the background is a pedestal spatially coextensive with the test (Nacer *et al.*, 1989) and when the test is of long duration (King-Smith and Carden, 1976).

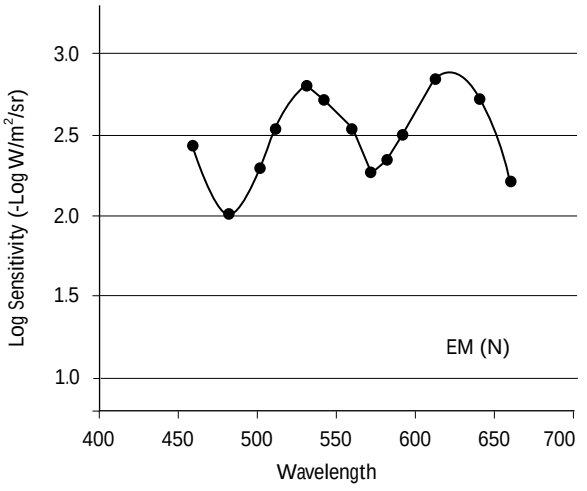


Figure 3.14 **Increment threshold spectral sensitivity function for 2° test stimuli presented on a coextensive 4600 K, 800 td pedestal within a larger 200 td surround. (Data replotted from Figure 6 of Miyahara *et al.*, 1996.)**

Interpretation: The characteristic lobes and notches of the increment spectral sensitivity on a white background are usually interpreted as reflecting spectrally opponent processing (see Chapter 6). Detection in the spectrally opponent channels is most comparable to KC-pathway 'blue-on,' and PC-pathway 'green-on' and 'red-on' responses (Sperling and Harwerth, 1971; Thornton and Pugh, 1983). The channels are weighted and fit to the data. This approach does not specify adaptation in a mechanistic form, but the weights serve this purpose. An alternative interpretation is in terms of psychophysically specified achromatic and chromatic visual pathways (King-Smith and Carden, 1976). The white background is considered a powerful adapting stimulus for the achromatic but not the two chromatic pathways. The lobes reflect weighted chromatic opponent activity; the notches represent weighted achromatic activity. This approach also does not specify adaptation in a mechanistic form.

3.4 CHROMATIC DISCRIMINATION

Discrimination refers to the ability to detect a difference between two lights that differ on some physical continuum. In studies of chromatic discrimination, the lights may differ in wavelength, in colorimetric purity, or in chromaticity, but do not differ in luminance. The threshold represents some criterion change presumably correlated with a difference in a neural response to the two lights. Thresholds may be expressed in wavelength steps or in chromaticity steps. A modern approach that attempts to relate detection and discrimination uses cone troland units (defined in section 3.2), which allows comparison of detection and discrimination in terms of quantal excitation of the receptors. If a surround is present and discrimination is measured from the surround chromaticity, the threshold may conceptually be considered as a detection, and related to the detection experiments described in

section 3.3. In modern studies of chromatic discrimination, the test chromaticity may differ from the surround chromaticity.

3.4.1 HISTORICAL APPROACHES

Wavelength discrimination: Wavelength discrimination refers to the ability of an observer to detect chromatic differences along the spectrum locus. In the classical wavelength discrimination experiment, the observer views a bipartite field, one half filled with light of a standard wavelength and the other with light of a comparison wavelength. Both standard and comparison fields are narrow spectral bands of light that are varied in spectral composition and radiance.

A common procedure is the step-by-step method in which both fields are initially of identical spectral composition and radiance (isomeric fields). The wavelength of the comparison field is changed in small steps (one nanometer or less) and the observer adjusts the radiance of the comparison field following each change to seek a match. Discrimination threshold is reached when the observer reports that the fields do not appear identical regardless of the radiance of the comparison field. The wavelength discrimination step is expressed in terms of the difference in wavelength, $\Delta\lambda$, between the standard and comparison fields. The procedure is repeated for a series of standard wavelengths throughout the visible spectrum. Data are reported either as the $\Delta\lambda$ in one direction, for example scaling toward longer wavelengths, or as the average of $\Delta\lambda$ for comparison lights scaled in both directions.

In an alternative technique, the standard deviation of a repeated series of color matches is taken as being proportional to the discrimination step. In this procedure, the two fields are initially different in wavelength. The observer adjusts the wavelength and radiance of the comparison field until it appears to be the same as the standard. The procedure is repeated many times, in order to determine the standard deviation of the comparison field settings. In terms of experimental convenience, the step-by-step method is more rapid but it requires absolute calibration of wavelength for both the standard and the comparison fields. The standard deviation procedure is time-consuming but may be more accurate as it shows less dependence on criterion changes of the observer. This technique also bridges measurements of color matching and color discrimination. The standard deviation of a set of color matches can be used as an index of color discrimination.

Figure 3.15 compares wavelength discrimination data from several laboratories. MacAdam's (1942) standard deviation data appear to be approximately one-fifth the discrimination thresholds measured by step-by-step procedures. The peaks and valleys vary somewhat among the different authors. A similar variation occurs among the functions measured in different individuals.

Colorimetric purity discrimination: Colorimetric purity discrimination typically refers to measurements of the least colorimetric purity, p_c, the minimum amount of spectral light that allows a mixture of a spectral light and white to be distinguished from white.

$$p_c = L_\lambda/(L_w + L_\lambda) \qquad (3.69)$$

where L_λ is the luminance of the spectral color and Lw is the luminance of the white. Figure 3.16(A) shows the data expressed as the reciprocal (p_c) of least colorimetric purity as a function of retinal illuminance. The results show a minimum in the 570–580 nm region, and best discrimination at the spectral extremes.

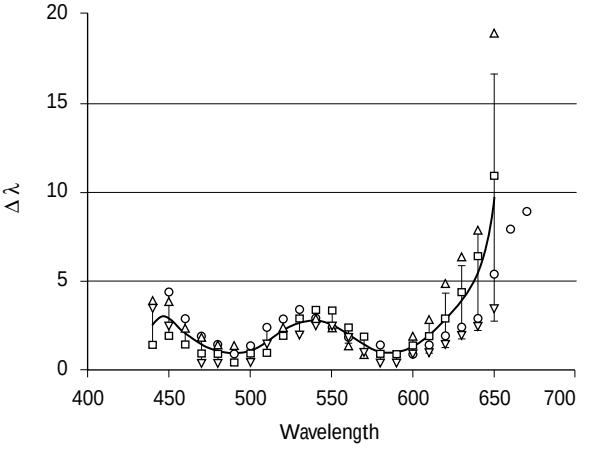


Figure 3.15 **Wavelength discrimination plotted as a function of wavelength. The symbols show data from four observers (Pokorny and Smith, 1970). The solid line is the average.**

Colorimetric purity of a sample, S, denoted by its chromaticity coefficients (x_S, y_S) in the CIE diagram is related to excitation purity, pe (section 3.2) by:

$$p_c = (y_\lambda / y_S) p_e \qquad (3.70)$$

where y_λ is the chromaticity coefficient of the dominant wavelength of sample S and y_S is the chromaticity coefficient of the sample.

If colorimetric purity discrimination is measured as the first step from the spectrum (i.e. least amount of white added to a spectral light), the results in the literature are rather variable but show a much flatter function than when colorimetric purity is measured as the first step from white (Jones and Lowry, 1926; Martin *et al.*, 1933; Wright and Pitt, 1937; Kaiser *et al.*, 1976). Yeh, Smith, and Pokorny (1993) reexamined this question carefully and found the shape of the first step from the spectrum function to be highly retinal illuminance dependent with flatter functions at higher light levels (Figure 3.16B).

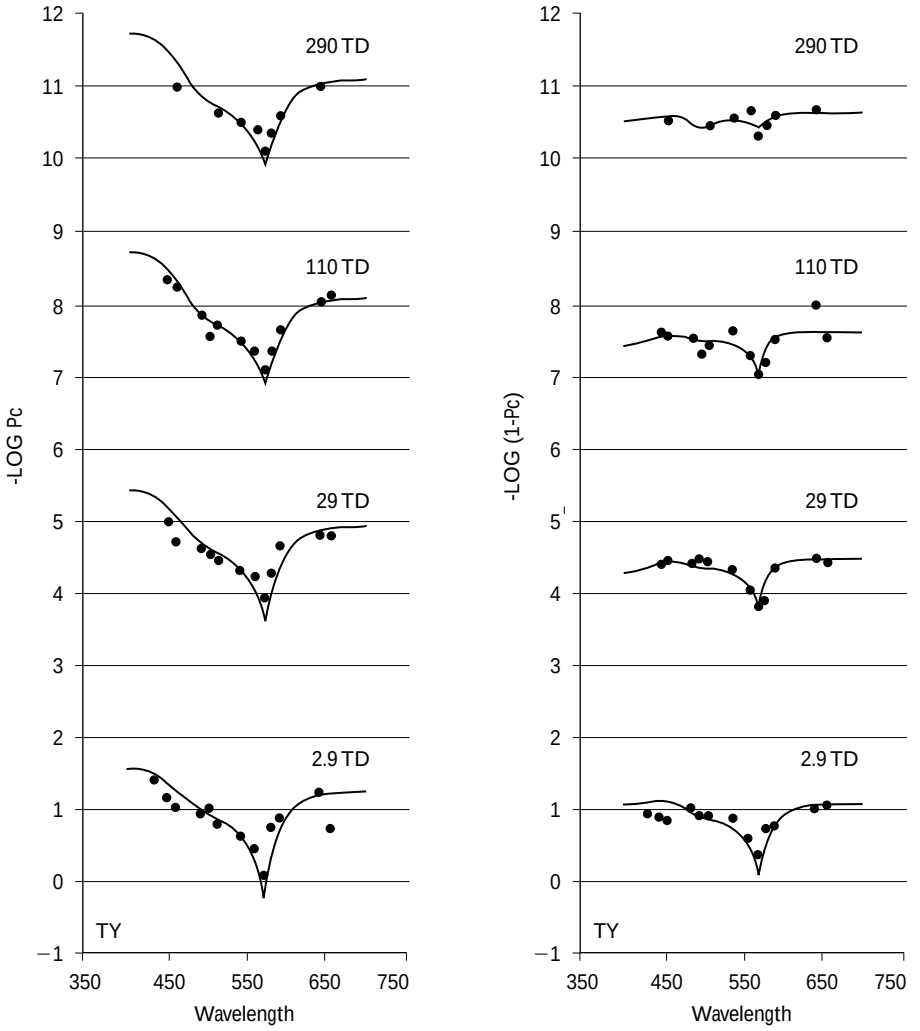


Figure 3.16 **Colorimetric purity discrimination plotted as a function of wavelength. Left panel shows colorimetric purity thresholds measured from white and right panel shows colorimetric purity thresholds measured from the spectrum. Data for the higher luminance levels are successively scaled vertically by 3 log units. Data are shown for one observer, data of four other observers showed closely similar trends. (From Yeh *et al.*, 1993.)**

MacAdam ellipses: Wavelength and colorimetric purity discrimination represent two special cases of chromaticity discrimination: discriminations along the spectrum locus and discrimination along axes between white and the spectrum locus. It is possible to sample chromatic discrimination systematically starting at an arbitrary chromaticity (e.g. Wright, 1941). The data are represented in the 1931 CIE chromaticity diagram (MacAdam, 1942; Brown and MacAdam, 1949; Wyszecki and Fielder, 1971). MacAdam used the standard deviation of color matches to represent chromaticity discrimination. For a number of different points in chromaticity space, MacAdam derived a series of discrimination ellipses, which represent the discriminable distance for a number of directions from each point. MacAdam ellipses represent data from a single observer. Figure 3.17 shows MacAdam's data plotted in the 1931 chromaticity diagram with each of the ellipses representing ten times the measured standard deviations.

In the equiluminant plane, discriminations are based solely on chromaticity differences. It is also possible to evaluate the joint effects of chromaticity and luminance in determining discrimination steps (Brown and MacAdam, 1949; Noorlander *et al.*, 1980). These data also describe ellipsoids in a three-dimensional chromaticity and luminance space (Poirson and Wandell, 1990).

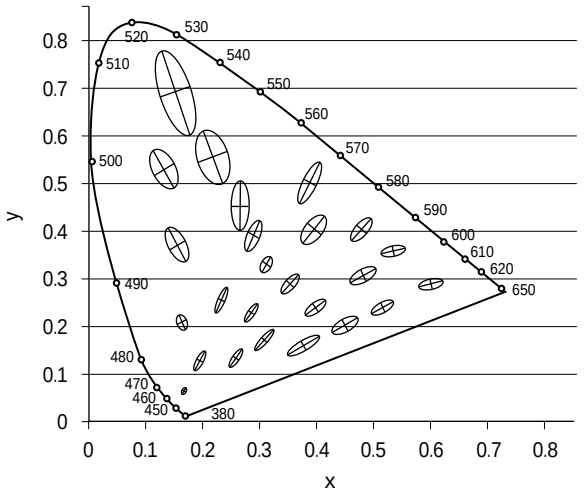


Figure 3.17 **The MacAdam ellipses plotted in the CIE chromaticity diagram. (From MacAdam, 1942.)**

Watson (1911) and Tyndall (1933) investigated the effects of adding white light to discrimination fields composed of spectral lights. For wavelengths greater than 490 nm, the discrimination step is increased with increases in the white light content. Tyndall extended these observations into the short wavelength region of the spectrum and found rather striking results. For 455 nm fields, discrimination improved with the addition of white light. Discrimination improved and was optimal when the white light was four to ten times higher in luminance than the spectral discrimination lights. This improvement has been termed the Tyndall effect. Polden and Mollon (1980) coined the term 'combinative euchromatopsia' to describe the enhanced sensitivity to hue differences.

3.4.2 EXPERIMENTAL VARIABLES

The effects of the luminance level, spatial structure, temporal presentation and retinal position are important in chromatic discrimination. The effect of surrounds is considered separately in section 3.4.4.

Effect of field size: Chromatic discrimination improves when the field size is increased from 2° to 10° by a factor of about two (Brown, 1952; Wyszecki and Stiles, 1982). The improvement is independent of the sampling direction in chromaticity space. Chromatic discrimination deteriorates when the field size is reduced below 1° (Pokorny and Smith, 1976). Chromatic discriminations are possible for bipartite fields as small as 3′ (MacAdam, 1959). Steady fixation of small fields (20′ or less) leads to a deterioration of discrimination mediated by the SWS cones (small field tritanopia, König, 1894; Thomson and Wright, 1947). The small field tritanopic effect also occurs with steady fixation of parafoveal fields (Hartridge, 1945; Thomson and Wright, 1947) and is reduced or eliminated by employing a scanning or glance technique (Bedford and Wyszecki, 1958b; McCree, 1960).

Effect of retinal illuminance: Discrimination functions show little change over a range of retinal illuminance of 100–3500 troland (Bedford and Wyszecki, 1958a; Cornu and Harlay, 1969). However, discrimination deteriorates at low

levels of retinal illuminance. The effect is particularly marked for discriminations based upon SWS cones (Brown, 1951; Verriest *et al.*, 1963; Knoblauch *et al.*, 1987). There is a strong interaction between field size and luminance; the deterioration of SWS cone discrimination is more pronounced with the joint reduction of field size and luminance (Farnsworth, 1955; Clarke, 1967; Yonemura and Kasuya, 1969).

The gap effect: Discrimination is also affected by the presence of a separation between two halves of a bipartite field (Sharpe and Wyszecki, 1976; Boynton *et al.*, 1977). The effect of introducing a gap is not the same for different types of discriminations. Luminance discrimination is best when the two half fields are precisely juxtaposed. Chromatic discrimination for red–green discriminations is unimpaired by a gap and discriminations based on differential SWS cone excitations improve.

Temporal presentation: Wavelength discrimination improves with increasing exposure duration but there is little agreement in the literature as to the optimal exposure (Farnsworth, 1961; Siegel, 1965; Regan and Tyler, 1971; Hita *et al.*, 1982). Discrimination deteriorates if stimuli are presented successively rather than simultaneously (Uchikawa and Ikeda, 1981; Uchikawa, 1983; Sachtler and Zaidi, 1992; Jin and Shevell, 1996).

Retinal location: Peripheral fixation of the fields results in a marked reduction of discrimination, particularly in the 490–530 nm (Moreland, 1972). Rods have been implicated as the source of degradation of discrimination (Lythgoe, 1931; Stabell and Stabell, 1977; Stabell and Stabell, 1982).

Temporal and spatial contrast sensitivity: A number of studies have measured the contrast sensitivity function for equiluminous chromatic alternation in space (van der Horst *et al.*,1967; Hilz and Cavonius, 1970), time (de Lange, 1958; Kelly and van Norren, 1977; Wisowaty, 1981; Swanson *et al.*, 1987), and joint variation of space and time (van der Horst and Bouman, 1969; Noorlander *et al.*, 1981). The studies report that equiluminous chromatic modulation showed a low pass function with poor sensitivity to high spatial or temporal frequencies.

Assessment of the spatial contrast sensitivity for equiluminant stimuli is technically demanding due to the presence of chromatic aberration in the eye. Chromatic aberration can introduce unwanted luminance information in the nominally equiluminous grating. Since the data reveal that a pure color grating has low contrast sensitivity, chromatic aberration must be carefully minimized (Mullen, 1985) or eliminated (Sekiguchi *et al.*, 1993).

A number of techniques have been used to assess the temporal and spatial characteristics of isolated receptor mechanisms. In general, the LWS and MWS mechanism have similar spatial (Brindley, 1954; Green, 1968; Cavonius and Estévez, 1975) and temporal properties (Brindley *et al.*, 1966; Green, 1969; Estévez and Cavonius, 1975). The SWS mechanism exhibits lower spatial (Stiles, 1949; Brindley, 1954; Green, 1968) and temporal (Brindley *et al.*, 1966; Green, 1969; Kelly, 1974; Wisowaty and Boynton, 1980) resolution.

3.4.3 MODERN APPROACH TO CHROMATICITY DISCRIMINATION

The Boynton and Kambe experiment: Boynton and Kambe (1980) performed a systematic sampling of chromaticity space in a constant luminance plane. Their study showed notable differences from the earlier studies. One major difference was that discrimination was measured along theoretically critical axes in the equiluminant plane which correspond to the axes of the MacLeod–Boynton chromaticity space (section 3.2.5). One set of axes maintained a constant amount of SWS cone stimulation, s/(l+m) and discrimination was measured for test chromaticities varying in their ratios of LWS to MWS cone stimulation, l/(l+m). A second set of axes maintained a constant LWS to MWS cone stimulation l/(l+m) and discrimination was evaluated for test chromaticities varying in SWS cone excitation, s/(l+m). The l/(l+m) cone discriminations showed a minimum. The minimum occurred near a value of 0.667 for l/(l+m) which is comparable to data obtained with a surround chromaticity at the equal energy spectrum (see section 3.4.4). The l/(l+m) discriminations were

primarily independent of the level of s/(l+m), although there appeared a small contribution of SWS cone activity at high s/(l+m) levels. On the S cone axis, discrimination was dependent on s/(l+m) at the test chromaticity. Further, s/(l+m) axis discriminations were affected only by the level of s/(l+m) and were independent of the l/(l+m) component of the stimulus. Subsequently, Krauskopf, Williams and Healy (1982) confirmed the independence of the l/(l+m) and s/(l+m) lines in an experiment using chromatic adaptation. They found highly selective chromatic effects with an adaptation field that was modulated in a l/(l+m) direction. The l/(l+m) discrimination was impaired by previous adaptation to a l/(l+m) stimulus, but discrimination along the s/(l+m) line was unaffected. Conversely, adaptation on a s/(l+m) affected s/(l+m) discrimination but not l/(l+m) discriminations. Luminance modulation had little effect on chromatic thresholds. Chromatic modulation had little effect on luminance thresholds.

The Boynton and Kambe data when expressed in chromaticity showed general agreement with MacAdam's (1942) ellipses. The Boynton and Kambe experiment was designed to yield high criterion color-difference steps. The observer was required to identify the color direction at threshold. Boynton and Kambe's discrimination steps are equivalent to approximately 13 of MacAdam's standard deviations.

A second major advance in the Boynton and Kambe formulation lies in the ability to relate discrimination data to detection by use of the cone troland. Boynton and Kambe showed that at 115 td optimal discrimination on the l/(l+m) axis required an increase of one L td (accompanied by a decrease of one M td). In comparison, optimal discrimination on the s/(l+m) axis required an increase of eight S td.

3.4.4 THE EFFECT OF SURROUNDS

Surrounds are important since they control the state of adaptation, as was clear from detection experiments. In early studies it was established that chromatic discrimination is best when the test chromaticity is at or near the chromaticity of the surround (Brown, 1952; Hurvich and Jameson, 1961; Pointer, 1974; Loomis and Berger, 1979). Modern studies, assessing discrimination on LWS, MWS cone, and on SWS cone excitation axes have confirmed and extended this finding (Krauskopf *et al.*, 1982; Zaidi *et al.*, 1992; Miyahara *et al.*, 1993; Smith *et al.*, 2000). The data show that l/(l+m) cone discriminations show a symmetrical V shape with minimum near the l/(l+m) chromaticity of the surround (Figure 3.18). Chromatic discrimination data obtained with no surround show a shallower V shape whose minimum coincides with that obtained when the surround chromaticity is that of the equal energy spectrum (Boynton and Kambe, 1980; Yeh *et al.*, 1993). Discriminations on the s/(l+m) axis also show a V shape with a minimum near the s/(l+m) cone chromaticity of the surround. The V is not symmetrical (Zaidi *et al.*, 1992; Miyahara *et al.*, 1993), rising more sharply when the s/(l+m) level of the test chromaticity is higher than that of the surround (Figure 3.19).

3.4.5 INTERPRETATION

An interpretation of chromatic discrimination data at equiluminance can be viewed in terms of spectral signals generated in the PC-pathway. A PC-pathway retinal ganglion cell has a steady resting level to a steady EES adapting field of 100–1000 trolands. If the field luminance or

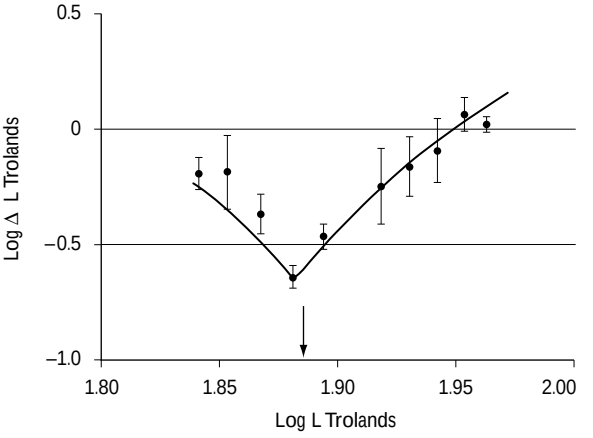


Figure 3.18 **Chromatic discrimination in the equiluminant plane for lights varying only in their l/(l+m) content. The observer is adapted to the equal energy spectrum. The data are expressed in L tds (see section 3.1) with log δL plotted vs. log L at the starting chromaticity. (From unpublished data of the authors.)**

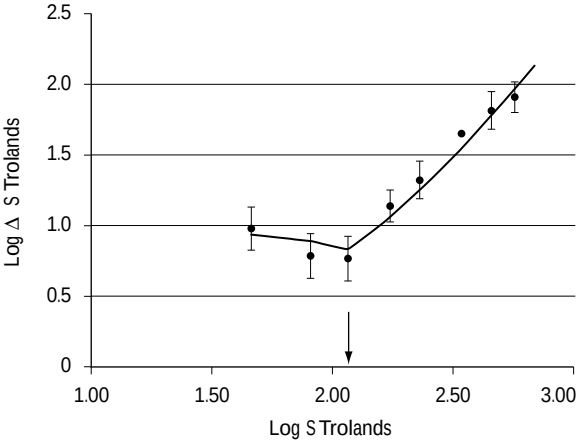


Figure 3.19 **Chromatic discrimination in the equiluminant plane for lights varying only in their s/(l+m) content. The observer is adapted to the equal energy spectrum. The data are expressed in S tds (see section 3.1) with log δS plotted vs. log S at the starting chromaticity. (From unpublished data of the authors.)**

chromaticity is changed, there is a return to near the resting level. This is indicative of almost complete adaptation to both chromaticity and luminance, occurring before or at the generation of the PC-pathway retinal ganglion cell signal. This adaptation could be either multiplicative or subtractive, as discussed in section 3.3. If a brief chromatic pulse is presented, the cell responds, showing a Naka–Rushton saturation function (Figure 3.12 in section 3.3) as the pulse is increased from the adaptation state. The V-shapes of chromatic discrimination can be viewed as characteristic of saturation functions following an earlier stage of neural adaptation (Zaidi *et al.*, 1992). At the point of the V, discrimination is assessed at the adapting chromaticity. These discriminations can also be considered as detections (section 3.3). The decrease in discrimination ability at low light levels is consistent with the idea that the neural adaptation becomes effective only above 1–10 trolands in the P pathway.

3.5 CONGENITAL COLOR DEFECT

Congenital color defects represent a hereditary, stationary condition in which there is an abnormality of color matching and/or color discrimination. Other visual function, including visual acuity, is normal. Such defects represent only a subset of the classification of human color defect which includes color vision problems accompanying hereditary and acquired eye disorders (Pokorny *et al.*, 1979). Congenital color defects have fascinated visual science since their discovery at the end of the eighteenth century. The fascination lies in the idea that color defects represent 'mistakes' of nature that can help elucidate the mechanisms of normal color vision. Today it is recognized that congenital color defects arise because of point mutations, rearrangements, and deletions of the opsin genes that determine the structure and function of the cone visual photopigments.

Congenital color defects may be classified on two dimensions: a qualitative dimension which states how color vision is affected and a quantitative dimension which describes the extent of discrimination loss. There are three major qualitative categories of defect, termed protan, deutan, and tritan.[3] The most common congenital defects are the protan and deutan defects. These defects show X-chromosome-linked inheritance and occur in 8% of the European male and 0.4% of the European female population (4–5% of the total European white population). The incidence in the United States is considered to be higher (Paulson, 1973). The tritan defect is rarer, occurring equally in both sexes in about 1 in 15 000 to 1 in 50 000 (0.002–0.007%) of the European population. The tritan defect has autosomal dominant inheritance. The inheritance, incidence, and classification of these defects are summarized in Table 3.5.

3.5.1 THE PROTAN AND DEUTAN DEFECTS

There are two qualitatively different forms of X-chromosome linked congenital defect. Protan observers show spectral sensitivity with long wavelength luminosity loss. Deutan observers show spectral sensitivity in the normal range at long wavelengths. The protan and deutan defects include a dichromatic form called protanopia and deuteranopia respectively. In the dichromatic defect, the affected observer requires only two primaries for full spectrum color matching. Protanopia and deuteranopia

Table 3.5 The inheritance, incidence and classification of congenital color vision defects

Type	Inheritance	Incidence	Classification
Protan and Deutan	X-chromosome linked recessive	8–10%(males) <1%(females)	
Protanopia		1%(males)	Dichromatic
Deuteranopia		1%(males)	Dichromatic
Protanomaly		1%(males)	Trichromatic
Deuteranomaly		5%(males)	Trichromatic
Tritan defect	Autosomal dominant	0.002–0.007	Dichromatic and Trichromatic

have traditionally been regarded as a 'reduction' form of color vision (von Kries, 1897). The dichromat was considered to lack function of one of the normal color fundamentals. The spectral sensitivity of protanopes and deuteranopes is shown in Figure 3.20. There are also trichromatic forms called protanomalous trichromacy (protanomaly) or deuteranomalous trichromacy (deuteranomaly). The color matches of protanomalous and deuteranomalous trichromats differ from each other and from those of normal trichromats. Anomalous trichromacy was historically regarded as an 'alteration' system (von Kries, 1897), but modern interpretation is in terms of inheritance of polymorphic forms of either the LWS (deuteranomalous) or the MWS (protanomalous) cone photopigments (see section 3.2.6).

The severity in discrimination loss of the protan and deutan defects varies considerably. The dichromatic form is more severe than the trichromatic form. It should be noted that variation is minimal within a family pedigree, i.e. both the qualitative and the quantitative extent of the defect are inherited. Color confusions and spectral sensitivity of protanopes and protanomalous trichromats or of deuteranopes and deuteranomalous trichromats are qualitatively similar, justifying the inclusive terms 'protan' and 'deutan' (Farnsworth, 1947). These similarities also allow the design of rapid screening tests for X-chromosome linked defects.

Full-spectrum colorimetry, spectral sensitivity, and wavelength and colorimetric purity discrimination functions for a limited number of observers have been described in the literature (Pokorny *et al.*, 1979). Some major features of protan and deutan color defects are summarized in Table 3.6.

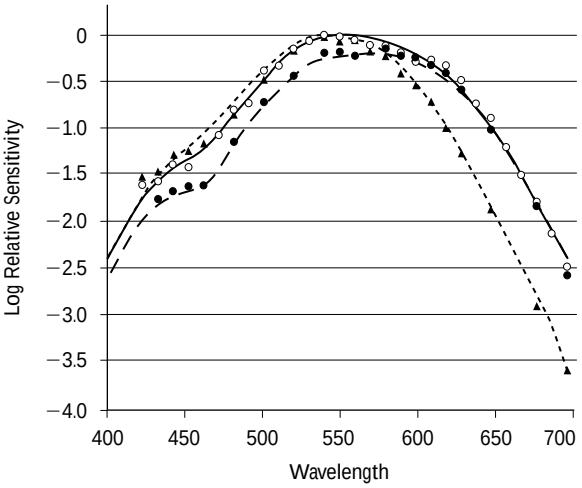


Figure 3.20 **Dichromatic spectral sensitivity data. The symbols show the threshold data of Hsia and Graham. The solid lines represent the Smith and Pokorny (1975) fundamentals of Figure 3.7, adjusted vertically by eye to pass near the data.**

The assumption that protanopes and deuteranopes lack one of the normal fundamentals allows the dichromatic color matches to be represented in the normal chromaticity diagram. The dichromatic color matches are normalized in the same manner as the normal matches. The dichromatic chromaticity coordinates fall on a line joining the two primaries. The chromaticity coordinates are joined to the corresponding test wavelength, forming a confusion line. The data reveal axes of discrimination loss characteristic of the type of defect. In the Judd revised chromaticity diagram, the confusion lines converge at a point, called the copunctal point (Figure 3.21). The copunctal point is considered the locus of the 'missing fundamental.' It is these copunctal points that form the new primaries in

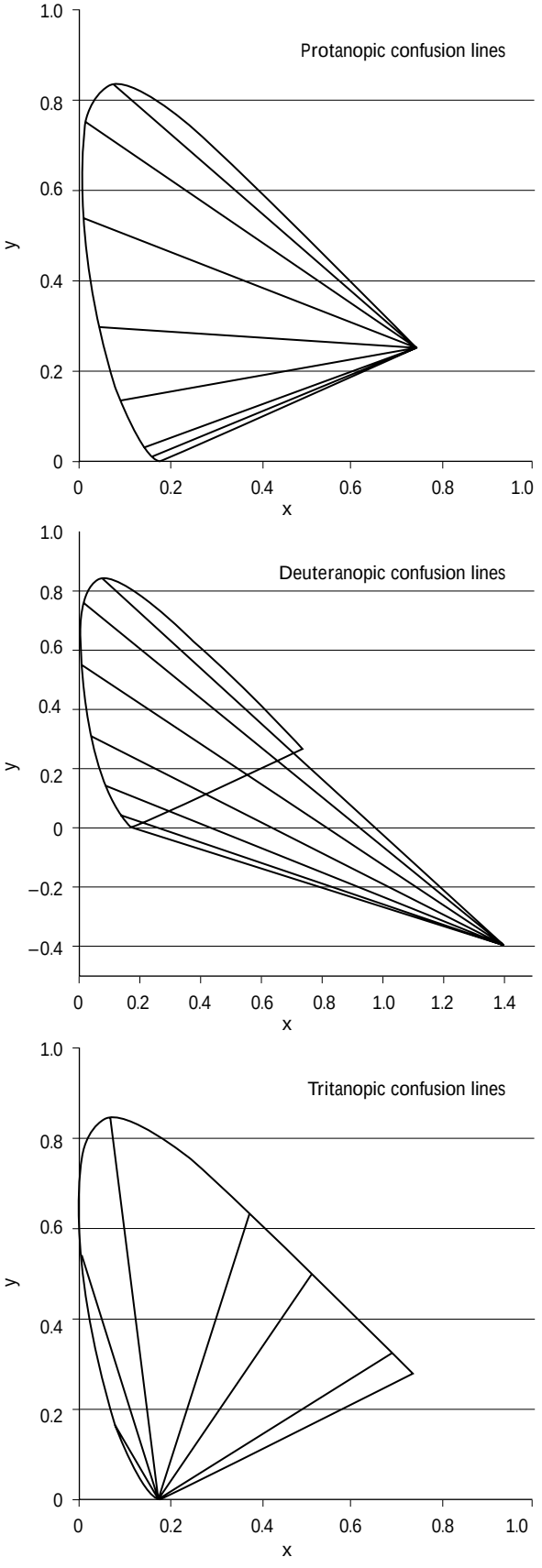
Table 3.6 Major features of congenital color vision defects

Characteristic	Normal	Protan	Deutan	Tritan
Number of primaries used in color mixture	3	2(P) 3(PA)	2(D) 3(DA)	2 or 3
Neutral point (dichromats only)		494 nm	499 nm	570 nm
Copunctal point (dichromats only)		$X_p = 0.7635$ $Y_p = 0.2365$	$X_d = 1.4000$ $Y_d = -0.400$	$X_t = 0.1748$ $Y_t = 0.0000$
λ_{max} of luminosity	555 nm	540 nm	560 nm	555 nm
Minimum delta λ (best wavelength discrimination)	590 nm	490 nm	495 nm	590 nm
Minimum Pc (worst purity discrimination)	570 nm	494 nm	499 nm	570 nm

transforming from the Judd color matching functions to a set of König fundamentals (see section 3.2.5). One confusion line passes from the chromaticity coordinates through the equal energy spectrum to the spectrum locus. The spectral wavelength at the intersection is called the neutral point.

The X-chromosome linked defects are diagnosed and classified using a specialized color match, the Rayleigh equation. The Rayleigh equation presents a match of a 589 nm spectral test light to a mixture of two spectral primaries, 545 nm and 670 nm. Since the primaries fall near the linear portion of the long wavelength spectrum locus, a third primary is unnecessary. An anomaloscope is used to measure the Rayleigh match; the instrument is designed so that the mixture field has fixed energy levels, only the proportion of 545 nm:670 nm lights is varied. The 589 nm test is variable in luminance. For some primary ratios, the trichromatic observer can adjust the 589 nm test field and obtain a match. Protanopic and deuteranopic observers can match all the primary ratios with suitable adjustment of the 589 nm test luminance. Normal, protanomalous and deuteranomalous trichromats use very different 545 nm:670 nm primary ratios in their matches, thus allowing classification with high specificity. The accepted classification of X-chromosome linked defects, based on Rayleigh matching with a 2° field of low photopic luminance is shown in Table 3.5. This classification has been extended (Franceschetti, 1928) and replicated in large-scale studies (Schmidt, 1955; Helve, 1972). It is important to recognize that change in the parameters of study (e.g. test field size, primaries, etc.) will not yield the same classification. For example, few dichromats will accept the full range of primary ratio with an 8° test field. Most would then be classified as anomalous trichromats. Screening tests based upon discrimination rather than color matching do not permit conclusive classification of color defect nor usually do they indicate the severity of discrimination loss.

The genetic study of the protan and deutan defects has revealed point mutations (single nucleotide changes) or deletions within the opsin genes on the X chromosome. The opsins for normal MWS and LWS photopigments lie in a tandem array on the X chromosome. There may be multiple copies of these genes and the factors controlling their expression are not yet delineated. The nucleotide sequences of the MWS and LWS opsin genes are very similar. Only a few nucleotide changes differentiate whether the absorption spectrum will be that of an LWS or MWS photopigment. Some variation in these nucleotides (polymorphism) occurs naturally in the color-normal population. The protan and deutan defects are correlated with



more pronounced alterations or with deletions of the opsin genes on the X-chromosome. The qualitative classification is well correlated with the type of gene alteration. The quantitative classification is less well correlated with the gene array. The correlation of genotype and phenotype remains an area of concentrated research.

3.5.2 TRITAN DEFECTS

The tritan defect is characterized by a lack of function of the mechanism that allows normals to discriminate colors that differ by the amount of short-wavelength light they contain. Discriminations dependent on LWS and MWS cone function are normal. Tritans show discrimination loss for SWS-cone mediated discrimination but do not show a specific alteration in color matching that would implicate variation in the absorption spectrum of the SWS photopigment. Some characteristics of tritan color deficiency are summarized in Table 3.6.

There is considerable variability in chromatic discrimination both within and between family pedigrees with tritan defect. Classification by color matching is less clear-cut. Some tritans make dichromatic color matches, but the majority do not. Discrimination improves with increase in field size (Pokorny *et al.*, 1981). Wright (1952) used a 1.2° field to classify and define tritanopia. For dichromatic tritans, the color matches can be treated as for protanopes and deuteranopes. The characteristic confusion axes for tritan defect are shown in Figure 3.21. There is no analog of anomalous trichromacy in autosomal dominant tritan defect size (Pokorny *et al.*, 1981).

The genetic study of the tritan defect has revealed point mutations (single nucleotide changes) or deletions within an opsin gene on the 7th chromosome ascribed to the SWS photopigment. The type of mutation is consistent within a pedigree of tritan defect.

Figure 3.21 **Dichromatic confusion lines plotted on the Judd revised chromaticity diagram. The upper panel shows protanopic confusion lines, the middle panel shows deuteranopic confusion lines and the lower panel shows tritanopic confusion lines. (Copunctal points from Smith and Pokorny, 1975).**

3.6 ACKNOWLEDGMENT

Preparation work for this chapter was supported in part by NIH Grant EY00901.

3.7 NOTES

1 The CIE has recommended standard sources for use in colorimetric specification. See Wyszecki and Stiles (1982) for further information. Those mentioned in this chapter include Standard Illuminants A, B, and C. CIE Standard Illuminant A is an incandescent tungsten filament lamp operated at a color temperature of approximately 2854 K. CIE Standard Illuminant B approximates noon sunlight, with a correlated color temperature of approximately 4870 K. CIE Standard Illuminant C approximates an overcast skylight, with a correlated color temperature of approximately 6740 K (Wyszecki and Stiles, 1982).

2 Differences in these equations from those in Smith and Pokorny (1975) result from rounding errors in the original calculation. Dr. Yasuhiso Nakano kindly noted and corrected them.

3 In the older literature, these defects were termed 'red-blind,' 'green-blind,' and 'blue-blind.' These terms, implying failure to perceive whole domains of color percepts, were dropped for the more agnostic terms now used. Nonetheless, color terms such as 'red–green defect' for the protan and deutan defects and 'blue–yellow defect' for the tritan defect remain common. Most protans and deutans recognize a wide variety of colors in the real world.

3.8 REFERENCES

Abramov, I., Gordon, J., and Chan, H. (1991) Color appearance in the peripheral retina: effects of stimulus size. *Journal of the Optical Society of America A*, 8, 404–14.

Adelson, E.H. (1982) Saturation and adaptation of the rod system. *Vision Research*, 22, 1299–312.

Alpern, M. (1979) Lack of uniformity in colour matching. *Journal of Physiology (London)*, 288, 85–105.

Barlow, H.B. (1956) Retinal noise and absolute threshold. *Journal of the Optical Society of America*, 46, 634–9.

Barlow, H.B. (1957) Increment thresholds at low intensities considered as signal/noise discrimination. *Journal of Physiology*, 136, 469–88.

Barlow, H.B. (1972) Dark and light adaptation: psychophysics. In D. Jameson and L.M. Hurvich (eds.), *Handbook of Sensory Physiology, Visual Psychophysics* (Vol. VII/4, pp. 1–28). Berlin: Springer-Verlag.

Bedford, R.E. and Wyszecki, G.W. (1958a) Luminosity functions for various field sizes and levels of retinal illuminance. *Journal of the Optical Society of America*, 48, 406–11.

Bedford, R.E. and Wyszecki, G.W. (1958b) Wavelength discrimination for point sources. *Journal of the Optical Society of America*, 48, 129–35.

Bouma, P.J. (1947) *Physical Aspects of Colour*. Eindhoven: N.V. Philips Gloeilampenfabrieken.

Bouman, M.A. and Koenderink, J.J. (1972) Psychophysical basis of coincidence mechanisms in the human visual system. *Ergebnisse der Physiolgie*, 65, 126–72.

Bowmaker, J.K., Dartnall, H.J.A., Lythgoe, J.N., and Mollon, J.D. (1978) The visual pigments of rods and cones in the rhesus monkey, *Macaca mulatta*. *Journal of Physiology (London)*, 274, 329–48.

Boynton, R.M. (1963) Contributions of threshold measurements to color-discrimination theory. *Journal of the Optical Society of America*, 53, 165–78.

Boynton, R.M., Hayhoe, M.M., and MacLeod, D.I.A. (1977) The gap effect: chromatic and achromatic visual discrimination as affected by field separation. *Optica Acta*, 24, 159–77.

Boynton, R.M., Ikeda, M., and Stiles, W.S. (1964) Interactions among chromatic mechanisms as inferred from positive and negative increment thresholds. *Vision Research*, 4, 87.

Boynton, R.M. and Kambe, N. (1980) Chromatic difference steps of moderate size measured along theoretically critical axes. *Color Research and Application*, 5, 13–23.

Brindley, G.S. (1953) The effects on colour vision of adaptation to very bright lights. *Journal of Physiology (London)*, 122, 332–50.

Brindley, G.S. (1954) The summation areas of human colour-receptive mechanisms at increment threshold. *Journal of Physiology (London)*, 124, 400–8.

Brindley, G.S. (1970) *Physiology of the Retina and the Visual Pathway*, 2nd edn. Baltimore, MD: Williams and Wilkins.

Brindley, G.S., du Croz, J.J., and Rushton, W.A.H. (1966) The flicker fusion frequency of the blue–sensitive mechanism of colour vision. *Journal of Physiology (London)*, 183, 497–500.

Brown, W.R.J. (1951) The influence of luminance level on visual sensitivity to color differences. *Journal of the Optical Society of America*, 41, 684–8.

Brown, W.R.J. (1952) The effect of field size and chromatic surroundings on color discrimination. *Journal of the Optical Society of America*, 42, 837–44.

Brown, W.R.J. and MacAdam, D.L. (1949) Visual sensitivities to combined chromaticity and luminance differences. *Journal of the Optical Society of America*, 39, 808–34.

Burns, S. and Elsner, A. (1985) Color matching at high illuminances: the color-match-area-effect and photopigment bleaching. *Journal of the Optical Society of America A*, 2, 698–704.

Burns, S.A. and Elsner, A.E. (1993) Color matching at high illuminances: photopigment optical density and pupil entry. *Journal of the Optical Society of America A*, 10, 221–30.

Cavonius, C.R. and Estévez, O. (1975) Sensitivity of human color mechanisms to gratings and flicker. *Journal of the Optical Society of America*, 65, 966–8.

CIE (1926) *Proceedings*, 1924, (pp. 1–232). Cambridge: Cambridge University Press.

CIE (1964) *Proceedings*, 1963 (Vienna Session), Vol. B. (Committee Report E–1.4.1), Paris, Bureau Central de la CIE, pp. 209–20.

Clarke, F.J.J. (1967) Colour measurement in industry. The effect of field–element size on chromaticity discrimination. Paper presented at the Proceedings of a Symposium on Colour Measurement in Industry, London.

Cornsweet, T.N. (1970) *Visual Perception*. New York: Academic Press.

Cornu, L. and Harlay, F. (1969) Modifications de la discrimination chromatique en fonction de l'eclairement. *Vision Research*, 9, 1273–87.

Crawford, B.H. (1949) The scotopic visibility function. *Proceedings of the Physical Society, B*, 62, 321–34.

Crawford, B.H. (1965) Colour matching and adaptation. *Vision Research*, 5, 71–8.

Dartnall, H.J.A. (1957) *The Visual Pigments*. London: Methuen and Company.

de Lange, H. (1958) Research into the dynamic nature of the human fovea–cortex systems with intermittent and modulated light. *Journal of the Optical Society of America*, 48, 779–89.

Eisner, A., Fleming, S.A., Klein, M.L., and Mouldin, W.M. (1987) Sensitivities in older eyes with good acuity: Cross-sectional norms. *Investigative Ophthalmology and Visual Science*, 28, 1824–31.

Eisner, A. and MacLeod, D.I.A. (1981) Flicker photometric study of chromatic adaptation: selective suppression of cone inputs by colored backgrounds. *Journal of the Optical Society of America*, 71, 705–18.

Estévez, O. (1979) On the fundamental data base of normal and dichromatic vision. Doctoral Dissertation (Vol. 17). Amsterdam: Krips Repro Meppel.

Estévez, O. and Cavonius, C.R. (1975) Flicker sensitivity of the human red and green color mechanisms. *Vision Research*, 15, 879–81.

Farnsworth, D. (1947) *The Farnsworth Dichotomous Test for Color Blindness – Panel D–15*. New York: Psychological Corporation.

Farnsworth, D. (1955) Tritanomalous vision as a threshold function. *Die Farbe*, 4, 185–97.

Farnsworth, D. (1961) A temporal factor in colour discrimination. *NPL Symposium No. 8, Visual Problems of Colour* (Vol. 2, pp. 65–78). New York: Chemical Publishing Co.

Franceschetti, A. (1928) Die Bedeutung der Einstellungsbreite am Anomaloskop fur die Diagnose der eizelnen Typen der Farbensinnstorungen nebst Bemerkungen uber ihre Vererbungsmodus. *Schweizerische Medizinische Wochenschrift*, 52, 1273–8.

Fridrikh, L. (1957) Colour-combination curves for normal trichromats determined by direct energy measurements. *Biophysics* [translation of *Biofizika*], 2, 129–32.

Galbraith, W. and Marshall, P.N. (1985) A survey of transforms of the CIE 1931 chromaticity diagram with some new non-linear transforms and histological illustrations of their utility. *Acta Histochemica*, 77, 79–100.

Geisler, W.S. (1981) Effects of bleaching and backgrounds on the flash response of the cone system. *Journal of Physiology (London)*, 312, 413–34.

Goldstein, E.B. and Williams, T.P. (1966) Calculated effects of 'screening pigments', *Vision Research*, 6, 39–50.

Grassmann, H. (1853) Zur Theorie der Farbenmischung. *Annalen der Physik (Leipzig)*, 89, 60–84. [English trans. *Philosophical Magazine (London)*, 1854, 7: 254–64.]

Green, D.G. (1968) The contrast sensitivity of the colour mechanisms of the human eye. *Journal of Physiology (London)*, 196, 415–29.

Green, D.G. (1969) Sinusoidal flicker characteristics of the color-sensitive mechanisms of the eye. *Vision Research*, 9, 591–601.

Grigorovici, R. and Aricescu-Savopol, I. (1958) Luminosity and chromaticity in the mesopic range. *Journal of the Optical Society of America*, 48, 891–8.

Guild, J. (1931) The colorimetric properties of the spectrum. *Philosophical Transactions of the Royal Society London A*, 230, 149–87.

Hammond, B.R., Wooten, B.R., and Snodderly, D.M. (1997) Individual variations in the spatial profile of human macular pigment. *Journal of Optical Society of America, A*, 14, 1187–96.

Hartridge, H. (1945) The change from trichromatic to dichromatic vision in the human retina. *Nature*, 155, 657–62.

Hayhoe, M., Benimoff, N.I., and Hood, D.C. (1987) The time-course of multiplicative and subtractive adaptation processes. *Vision Research*, 27, 1981–96.

Hecht, S., Shlaer, S., and Pirenne, M.H. (1942) Energy, quanta and vision. *Journal of General Physiology*, 224, 665–99.

Helve, J. (1972) A comparative study of several diagnostic tests of colour vision used for measuring types and degrees of congenital red–green defects. *Acta Ophthalmologica Supplement*, 115, 1–64.

Hilz, R. and Cavonius, C.R. (1970) Wavelength discrimination measured with square-wave gratings. *Journal of the Optical Society of America*, 60, 273–7.

Hita, E., Romero, J., Jimenez del Barco, L., and Martinez, R. (1982) Temporal aspects of color discrimination. *Journal of the Optical Society of America*, 72, 578–82.

Hood, D.C. (1998) Lower-level visual processing and models of light adaptation. *Annual Review of Psychology*, 49, 503–35.

Hood, D.C. and Finkelstein, M.A. (1986) Sensitivity to light. In K.R. Boff, L. Kaufman, and J.P. Thomas (eds), *Handbook of Perception and Human Performance, Vol I: Sensory Processes and Perception*, pp. 5-1–5-66). New York: John Wiley and Sons.

Horner, R.G. and Purslow, E.T. (1947) Dependence of

anomaloscope matches on viewing distance or field size. *Nature*, 160, 23–4.
Hurvich, L.M. and Jameson, D. (1961) Opponent chromatic induction and wavelength discrimination. In R. Jung and H. Kornhuber (eds), *The Visual System: Neurophysiology and Psychophysics*. Berlin: Springer, pp. 144–52.
Ingling, C.R. and Martinez, E. (1981) Stiles' Π 5 mechanism: failure to show univariance is caused by opponent–channel input. *Journal of the Optical Society of America*, 71, 1134–7.
Jin, E.W. and Shevell, S.K. (1996) Color memory and color constancy. *Journal of the Optical Society of America A*, 13, 1981–91.
Jones, L.A. and Lowry, E.M. (1926) Retinal sensibility to saturation differences. *Journal of the Optical Society of America*, 13, 25–34.
Judd, D.B. (1930) Reduction of data on mixture of color stimuli. *Journal of Research National Bureau of Standards (USA)*, 4, 515–48.
Judd, D.B. (1935) A Maxwell triangle yielding uniform chromaticity scales. *Journal of the Optical Society of America*, 25, 24–35.
Judd, D.B. (1951a) Basic correlates of the visual stimulus. In S.S. Stevens (ed.), *Handbook of Experimental Psychology*. New York: Wiley, pp. 811–67.
Judd, D.B. (1951b) Colorimetry and artificial daylight. In Technical Committee No. 7 Report of Secretariat United States Commission, International Commission on Illumination, Twelfth Session, Stockholm, pp. 1–60.
Kaiser, P.K., Comerford, J.P., and Bodinger, D.M. (1976) Saturation of spectral lights. *Journal of the Optical Society of America*, 66, 818–26.
Kelly, D.H. (1974) Spatio-temporal frequency characteristics of color-vision mechanisms. *Journal of the Optical Society of America*, 64, 983–90.
Kelly, D.H. and van Norren, D. (1977) Two-band model of heterochromatic flicker. *Journal of the Optical Society of America*, 67, 1081–91.
King-Smith, P.E. and Carden, D. (1976) Luminance and opponent-color contributions to visual detection and adaptation and to temporal and spatial integration. *Journal of the Optical Society of America*, 66, 709–17.
King-Smith, P.E. and Webb, J.R. (1974) The use of photopic saturation in determining the fundamental spectral sensitivity curves. *Vision Research*, 14, 421–9.
Knoblauch, K., Saunders, F., Kusuda, M., Hynes, R., Podgor, M., Higgens, K.E., and deMosasterio, M. (1987) Age and illuminance effects in the Farnsworth–Munsell 100-hue test. *Applied Optics*, 26, 1441–8.
Koenderink, J.J., van de Grind, W.A., and Bouman, M.A. (1970) Models of retinal signal processing at high luminances. *Kybernetik*, 6, 227–37.
König, A. (1894) Uber den menschlichen Sehpurpur und seine Bedeutung fur das Sehen. *Akademie der Wissenschaften Berlin, Sitzungsberichte*, pp. 577–98.
König, A. and Dieterici, C. (1886) Die Grundempfindungen und ihre Intensitats–Vertheilung im Spectrum. *Akademie der Wissenschaften Berlin, Sitzungsberichte, Pt.2*, pp. 805–29.
König, A. and Dieterici, C. (1893) Die Grundempfindungen in normalen und anomalen Farben Systemen und ihre Intensitats-Verthielung im Spectrum. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 4, 241–347.
Krauskopf, J., Williams, D.R., and Heeley, D.W. (1982) Cardinal directions of color space. *Vision Research*, 22, 1123–31.
Le Grand, Y. (1968) *Light, Colour and Vision*, 2nd edn. London: Chapman and Hall, pp. 1–564.
Lee, B.B., Pokorny, J., Smith, V.C., Martin, P.R., and Valberg, A. (1990) Luminance and chromatic modulation sensitivity of macaque ganglion cells and human observers. *Journal of the Optical Society of America A*, 7, 2223–36.
Loomis, J.M. and Berger, T. (1979) Effects of chromatic adaptation on color discrimination and color appearance. *Vision Research*, 19, 891–901.
Lythgoe, R. (1931) Dark-adaptation and the peripheral colour sensations of normal subjects. *British Journal of Ophthalmology*, 15, 193–210.
MacAdam, D.L. (1942) Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America*, 32, 247–74.
MacAdam, D.L. (1959) Small-field chromaticity discrimination. *Journal of the Optical Society of America*, 49, 1143–6.
MacLeod, D.I.A. and Boynton, R.M. (1979) Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America*, 69, 1183–5.
Martin, L.C., Warburton, F.L., and Morgan, W.J. (1933) The determination of the sensitiveness of the eye to differences in the saturation of colours. *Great Britain Medical Research Council*, 188, 5–42.
Maxwell, J.C. (1860) On the theory of compound colors and relations of the colors of the spectrum. *Philosophical Transactions*, 150, 57–84. [Reprinted with commentary by Qasim Zaidi in: *Color Research and Application*, 1993, 18, 270–87.]
McCree, K.J. (1960) Small-field tritanopia and the effects of voluntary fixation. *Optica Acta*, 7, 317–23.
Merbs, S.L. and Nathans, J. (1992) Absorption spectra of human cone pigments. *Nature*, 356, 433–5.
Miles, W.R. (1954) Comparison of the functional and structural areas in human fovea. I. Method of entoptic plotting. *Journal of Neurophysiology*, 17, 22–38.
Miller, S.S. (1972) Psychophysical estimates of visual pigment densities in red–green dichromats. *Journal of Physiology (London)*, 223, 89–107.
Miyahara, E., Pokorny, J., and Smith, V.C. (1996) Increment threshold and purity discrimination spectral sensitivities of X-chromosome-linked color defective observers. *Vision Research*, 36, 1597–613.
Miyahara, E., Smith, V.C., and Pokorny, J. (1993) How surrounds affect chromaticity discrimination. *Journal of the Optical Society of America A*, 10, 545–53.

Mollon, J.D. and Polden, P.G. (1977) An anomaly in the response of the eye to light of short wavelengths. *Philosophical Transactions of the Royal Society of London – Series B: Biological Sciences*, 278, 207–40.

Moreland, J.D. (1972) Peripheral Colour Vision. In D. Jameson and L.M. Hurvich (eds), *Handbook of Sensory Physiology, Visual Psychophysics*, Vol. VII/4. Berlin: Springer-Verlag, pp. 517–36.

Moreland, J.D. and Bhatt, P. (1984) Retinal distribution of macular pigment. *Documenta Ophthalmologica Proceedings Series*, 39, 127–32.

Moreland, J.D. and Cruz, A. (1959) Colour perception with the peripheral retina. *Optica Acta*, 6, 117–51.

Moreland, J.D., Torczynski, E., and Tripathi, R. (1991) Rayleigh and Moreland matches in the ageing eye. *Documenta Ophthalmologica Proceedings Series*, 54, 347–52.

Mullen, K.T. (1985) The contrast sensitivity of human colour vision to red–green and blue–yellow chromatic gratings. *Journal of Physiology (London)*, 359, 381–400.

Nacer, A., Murray, I.J., and Carden, D. (1989) Interactions between luminance mechanisms and colour opponency. In C.M. Dickinson and I.J. Murray (eds), *Seeing Contour and Colour*. Oxford: Pergamon Press, pp. 357–60.

Neitz, M., Neitz, J., and Jacobs, G.H. (1991) Spectral tuning of pigments underlying red–green color vision. *Science*, 252, 971–3.

Noorlander, C., Heuts, M.J.G., and Koenderink, J.J. (1980) Influence of the target size on the detection threshold for luminance and chromaticity contrast. *Journal of the Optical Society of America*, 70, 1116–21.

Noorlander, C., Heuts, M.J.G., and Koenderink, J.J. (1981) Sensitivity to spatiotemporal combined luminance and chromaticity contrast. *Journal of the Optical Society of America*, 71, 453–9.

Palmer, D.A. (1978) Maxwell spot and additivity in tetrachromatic matches. *Journal of the Optical Society of America*, 68, 1501–5.

Palmer, D.A. (1981) Nonadditivity in color matches with four instrumental primaries. *Journal of the Optical Society of America*, 71, 966–9.

Paulson, H.M. (1973) Comparison of color vision tests used by the Armed Forces. *Color Vision*. National Academy of Sciences.

Pease, P.L., Adams, A.J., and Nuccio, E. (1987) Optical density of human macular pigment. *Vision Research*, 27, 705–10.

Pointer, M.R. (1974) Color discrimination as a function of observer adaptation. *Journal of the Optical Society of America*, 64, 750–9.

Poirson, A.B. and Wandell, B.A. (1990) The ellipsoidal representation of spectral sensitivity. *Vision Research*, 30, 647–52.

Pokorny, J. and Smith, V.C. (1970) Wavelength discrimination in the presence of added chromatic fields. *Journal of the Optical Society of America*, 69, 562–9.

Pokorny, J. and Smith, V. C. (1976) Effect of field size on red–green color mixture equations. *Journal of the Optical Society of America*, 66, 705–8.

Pokorny, J. and Smith, V.C. (1981) A variant of red–green color defect. *Vision Research*, 21, 311–17.

Pokorny, J. and Smith, V.C. (1986) Colorimetry and Color Discrimination. In K.R. Boff, L. Kaufman, and J P. Thomas (eds), *Handbook of Perception and Human Performance, Vol I: Sensory Processes and Perception*.) New York: John Wiley and Sons, pp. 8-1–8-51.

Pokorny, J. and Smith, V.C. (1997) How much light reaches the retina? In C.R. Cavonius (ed.), *Colour Vision Deficiencies XIII. Documenta Ophthalmologica Proceedings Series*, 59, 491–511.

Pokorny, J., Smith, V. C., and Lutze, M. (1987) Aging of the human lens. *Applied Optics*, 26, 1437–40.

Pokorny, J., Smith, V.C., and Starr, S.J. (1976) Variability of color mixture data – II. The effect of viewing field size on the unit coordinates. *Vision Research*, 16, 1095–8.

Pokorny, J., Smith, V.C., Verriest, G., and Pinckers, A.J.L.G. (eds) (1979) *Congenital and Acquired Color Vision Defects*. New York: Grune and Stratton.

Pokorny, J., Smith, V.C., and Went, L.N. (1981) Color matching in autosomal dominant tritan defect. *Journal of the Optical Society of America*, 71, 1327–34.

Polden, P.G. and Mollon, J.D. (1980) Reversed effects of adapting stimuli on visual sensitivity. *Proceedings of the Royal Society of London Series B*, 210, 235–72.

Polyak, S.L. (1957) *The Vertebrate Visual System*. Chicago: University of Chicago Press.

Pugh, E.N. (1976) The nature of the π-1 colour mechanism of W.S. Stiles. *Journal of Physiology (London)*, 257, 713–47.

Pugh, E.N.J. and Kirk, D.B. (1986) The pi mechanisms of WS Stiles: an historical review. [Review]. *Perception*, 15, 705–28.

Pugh, E.N.J. and Mollon, J.D. (1979) A theory of the π-1 and π-3 color mechanisms of Stiles. *Vision Research*, 19, 293–312.

Regan, D. and Tyler, C.W. (1971) Temporal summation and its limit for wavelength changes: an analog of Bloch's Law for color vision. *Journal of the Optical Society of America*, 61, 1414–21.

Richards, W. and Luria, S.M. (1964) Color-mixture functions at low luminance levels. *Vision Research*, 4, 281–313.

Ruddock, K.H. (1963) Evidence for macular pigmentation from colour matching data. *Vision Research*, 3, 417–29.

Rushton, W.A.H. (1972) Visual pigments in man. In H.J.A. Dartnall (ed.), *Handbook of Sensory Physiology*, Vol. VII/I. Berlin: Springer, pp. 364–94.

Sachtler, W.L. and Zaidi, Q. (1992) Chromatic and luminance signals in visual memory. *Journal of the Optical Society of America A*, 9, 877–94.

Sanocki, E., Shevell, S.K., and Winderickx, J. (1994) Serine/Alanine amino acid polymorphism of the L-cone photopigment assessed by dual Rayleigh-type color matches. *Vision Research*, 34, 377–82.

Schmidt, I. (1955) Some problems related to testing color vision with the Nagel anomaloscope. *Journal of the Optical Society of America*, 45, 514–22.

Schrödinger, E. (1920) Grundlinien einer Theorie der Farbenmetrik im Tagessehen. *Annalen der Physik (Leipzig)*, 63, 134–82. English translation in D.L. MacAdam (ed.), *Sources of Color Science*. Cambridge, MA: The MIT Press, 1970.

Schrödinger, E. (1925) Ueber das Verhaltnis der Vierfarben- zur Dreifarbentheorie. *Sitzungsberichte der Mathematisch–naturwissenschaftlichen Klasse der Kaiserlichen Akademie der Wissenschaften Wien, 134, Abt. IIa*, 471–90. [Commentary by Qasim Zaidi and National Translation Center translation 'On the relationship of four-color theory to three-color theory', *Color Research and Application*, 1994, 19, 37–47.]

Sekiguchi, N., Williams, D.R., and Brainard, D.H. (1993) Aberration-free measurements of the visibility of isoluminant gratings. *Journal of the Optical Society of America A*, 10, 2105–17.

Shapiro, A.G., Pokorny, J., and Smith, V.C. (1994) Rod contribution to large field color-matching. *Color Research and Application*, 19, 236–45.

Shapiro, A.G., Pokorny, J., and Smith, V.C. (1996) Cone-rod receptor spaces, with illustrations that use CRT phosphor and light-emitting-diode spectra. *Journal of the Optical Society of America A*, 13, 2319–28.

Sharpe, L.T. and Wyszecki, G. (1976) Proximity factor in color-difference evaluations. *Journal of the Optical Society of America*, 66, 40–9.

Shevell, S.K. (1977) Saturation in human cones. *Vision Research*, 17, 427–34.

Siegel, M.H. (1965) Color discrimination as a function of exposure time. *Journal of the Optical Society of America*, 55, 566–8.

Smith, V. and Pokorny, J. (1973) Psychophysical estimates of optical density in human cones. *Vision Research*, 13, 1099–202.

Smith, V.C. and Pokorny, J. (1975) Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research*, 15, 161–71.

Smith, V.C. and Pokorny, J. (1977) Large-field trichromacy in protanopes and deuteranopes. *Journal of the Optical Society of America*, 67, 213–20.

Smith, V.C. and Pokorny, J. (1996) The design and use of a cone chromaticity space. *Color Research and Application*, 21, 375–83.

Smith, V.C., Pokorny, J., Lee, B.B., and Dacey, D.M. (2001) Primate Horizontal Cell Dynamics: An Analysis of Sensitivity Regulation in the Outer Retina. *Journal of Neurophysiology*, 85, 545–58.

Smith, V.C., Pokorny, J., and Starr, S.J. (1976) Variability of color mixture data – I. Interobserver variability in the unit coordinates. *Vision Research*, 16, 1087–94.

Smith, V.C., Pokorny, J., and Sun, H. (2000) Chromatic contrast discrimination: Data and prediction for stimuli varying in L and M cone excitation. *Color Research and Application*, 25, 105–15.

Smith, V.C., Pokorny, J., and Zaidi, Q. (1983) How do sets of color-matching functions differ? In J.D. Mollon and L.T. Sharpe (eds), *Colour Vision: Physiology and Psychophysics*. London: Academic Press, pp. 93–105.

Speranskaya, N.I. (1959) Determination of spectrum color coordinates for twenty-seven normal observers. *Optics and Spectroscopy*, 7, 424–8.

Speranskaya, N.I. (1961) Methods of determination of the co-ordinates of spectrum colours. *NPL Symposium No. 8, Visual Problems of Colour* (Vol. 1, pp. 319–325). New York: Chemical Publishing Co.

Sperling, H.G. and Harwerth, R.S. (1971) Red–green cone interaction in the increment-threshold spectral sensitivity of primates. *Science*, 172, 180–4.

Stabell, B. and Stabell, U. (1976) Rod and cone contributions to peripheral colour vision. *Vision Research*, 16, 1099–104.

Stabell, U. and Stabell, B. (1977) Wavelength discrimination of peripheral cones and its change with rod intrusion. *Vision Research*, 17, 423–6.

Stabell, U. and Stabell, B. (1980) Variation in density of macular pigmentation and in short-wave cone sensitivity with eccentricity. *Journal of the Optical Society of America*, 70, 706–11.

Stabell, U. and Stabell, B. (1982) Color vision in the peripheral retina under photopic conditions. *Vision Research*, 22, 839–44.

Stiles, W.S. (1949) Increment thresholds and the mechanisms of colour vision. *Documenta Ophthalmologica*, 3, 138–63.

Stiles, W.S. (1955) 18th Thomas Young Oration. The basic data of colour-matching. *The Yearbook of the Physical Society*, pp. 44–65.

Stiles, W.S. (1959) Color vision: The approach through increment threshold sensitivity. *Proceedings National Academy of Sciences USA*, 45, 100–14.

Stiles, W.S. (1978) *Mechanisms of Colour Vision*. London: Academic Press.

Stiles, W.S. and Burch, J.M. (1955) Interim report to the Commission Internationale de l'Eclairage, Zurich, 1955, on the National Physical Laboratory's investigation of colour-matching. *Optica Acta*, 2, 168–81.

Stiles, W.S. and Burch, J.M. (1959) NPL colour-matching investigation: final report. *Optica Acta*, 6, 1–26.

Stockman, A., MacLeod, D.M., and Johnson, N.E. (1993) Spectral sensitivities of the human cones. *Journal of the Optical Society of America*, 10, 2491–521.

Stockman, A., MacLeod, D.M., and Vivien, J.A. (1993) Isolation of the middle- and long-wavelength sensitive cones in normal trichromats. *Journal of the Optical Society of America*, 10, 2471–90.

Stockman, A. and Mollon, J. (1986) The spectral sensitivities of the middle- and long-wavelength cones: an extension of the two-colour threshold technique of W.S. Stiles. *Perception*, 15, 729–54.

Stockman, A. and Sharpe, L.T. (2000) The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Research*, 40, 1711–37.

Swanson, W.H. and Fish, G.E. (1996) Age-related

changes in the color-match-area effect. *Vision Research*, 36, 2079–85.
Swanson, W.H., Ueno, T., Smith, V.C., and Pokorny, J. (1987) Temporal modulation sensitivity and pulse detection thresholds for chromatic and luminance perturbations. *Journal of the Optical Society of America A*, 4, 1992–2005.
Tan, K.E.W.P. (1971) *Vision in the Ultraviolet*. Utrecht: Drukkerij Elinkwijk.
Terstiege, H. (1967) Untersuchungen zum Persistenz- und Koeffizientensatz. *Die Farbe*, 16, 1–120.
Thomson, L.C. and Wright, W.D. (1947) The colour sensitivity of the retina within the central fovea of man. *Journal of Physiology (London)*, 105, 316–31.
Thornton, J.E. and Pugh, E.N.J. (1983) Red/green color opponency at detection threshold. *Science*, 219, 191–3.
Trezona, P.W. (1970) Rod participation in the 'blue' mechanism and its effect on colour matching. *Vision Research*, 10, 317–32.
Troland, L.T. (1915) The theory and practise of the artificial pupil. *Psychological Review*, 22, 167–76.
Troy, J.B. and Lee, B.B. (1994) Steady discharges of macaque retinal ganglion cells. *Visual Neuroscience*, 11, 111–18.
Troy, J.B. and Robson, J.G. (1992) Steady discharges of X and Y retinal ganglion cells of cat under photopic illuminance. *Visual Neuroscience*, 9, 535–53.
Tyndall, E.P.T. (1933) Chromaticity sensibility to wave-length difference as a function of purity. *Journal of the Optical Society of America*, 23, 15–24.
Uchikawa, K. (1983) Purity discrimination: successive vs simultaneous comparison method. *Vision Research*, 23, 53–8.
Uchikawa, K. and Ikeda, M. (1981) Temporal deterioration of wavelength discrimination with successive comparison method. *Vision Research*, 21, 591–5.
van der Horst, G.J.C. and Bouman, M.A. (1969) Spatiotemporal chromaticity discrimination. *Journal of the Optical Society of America*, 59, 1482–8.
van der Horst, G.J.C., de Weert, C.M.M., and Bouman, M.A. (1967) Transfer of spatial chromaticity contrast at threshold in the human eye. *Journal of the Optical Society of America*, 57, 1260–7.
van Norren, D. and Vos, J.J. (1974) Spectral transmission of the human ocular media. *Vision Research*, 14, 1237–44.
Verriest, G., Buyssens, A., and Vanderdonck, R. (1963) Etude quantitative de l'effet qu'exerce sur les resultats de quelques tests de la discrimination chromatique une diminution non selective du niveau d'un eclairage. *Revue d'Optique*, 3, 105–19.
Viénot, F. (1983) Can variation in macular pigment account for the variation of colour matches with retinal position? In J.D. Mollon and L.T. Sharpe (eds), *Colour Vision*. New York: Academic Press, pp. 107–16.
von Kries, J. (1897) Uber Farbensysteme. *Zeitschrift für Psychologie Physiologie Sinnesorg*, 13, 241–324.
Vos, J.J. (1972) Literature review of human macular absorption in the visible and its consequences for the cone receptor primaries (2F1972–17). TNO Report, Institute for Perception, Soesterberg The Netherlands.
Vos, J.J. (1978) Colorimetric and photometric properties of a 2° fundamental observer. *Color Research and Application*, 3, 125–8.
Vos, J.J. and Walraven, P.L. (1971) On the derivation of the foveal receptor primaries. *Vision Research*, 11, 799–818.
Wald, G. (1964) The receptors of human color vision. *Science*, 145, 1007.
Watson, W. (1911) Note on the sensibility of the eye to variations of wave-length. *Proceedings of the Royal Society (London)*, B84, 118–21.
Weale, R.A. (1988) Age and the transmittance of the human crystalline lens. *Journal of Physiology (London)*, 395, 577–87.
Werner, J.S., Donnelly, S.K., and Kliegl, R. (1987) Aging and human macular pigment density. *Vision Research*, 27, 257–68.
Wisowaty, J.J. (1981) Estimates for the temporal response characteristics of chromatic pathways. *Journal of the Optical Society of America*, 71, 970–7.
Wisowaty, J.J. and Boynton, R.M. (1980) Temporal modulation sensitivity of the blue mechanism: measurements made without chromatic adaptation. *Vision Research*, 20, 895–909.
Wright, W.D. (1929) A re-determination of the trichromatic coefficients of the spectral colours. *Transactions of the Optical Society*, 30, 141–64.
Wright, W.D. (1941) The sensitivity of the eye to small colour differences. *Proceedings of the Physical Society (London)*, 53, 93–112.
Wright, W.D. (1946) *Researches on Normal and Defective Colour Vision*. London: Henry Kimpton.
Wright, W.D. (1952) The characteristics of tritanopia. *Journal of the Optical Society of America*, 42, 509–20.
Wright, W.D. and Pitt, F.H.G. (1937) The saturation-discrimination of two trichromats. *Proceedings of the Physical Society (London)*, 49, 329–31.
Wyszecki, G. and Fielder, G.H. (1971) New color-matching ellipses. *Journal of the Optical Society of America*, 61, 1135–52.
Wyszecki, G. and Stiles, W.S. (1980) High-level trichromatic color matching and the pigment-bleaching hypothesis. *Vision Research*, 20, 23–37.
Wyszecki, G. and Stiles, W.S. (1982) *Color Science – Concepts and Methods, Quantitative Data and Formulae*, 2nd edn. New York: John Wiley and Sons.
Yeh, T., Pokorny, J., and Smith, V.C. (1993) Chromatic discrimination with variation in chromaticity and luminance: data and theory. *Vision Research*, 33, 1835–45.
Yeh, T., Smith, V.C., and Pokorny, J. (1989) The effect of background luminance on cone sensitivity functions. *Investigative Ophthalmology and Visual Science*, 30, 2077–86.
Yeh, T., Smith, V. C., and Pokorny, J. (1993) Colorimetric purity discrimination: data and theory. *Vision Research*, 33, 1847–57.

Yonemura, G.T. and Kasuya, M. (1969) Color discrimination under reduced angular subtense and luminance. *Journal of the Optical Society of America*, 59, 131–5.
Young, T. (1802) On the theory of light and colours. *Philisophical Transactions (London)*, 92, 12–48.
Zaidi, Q. (1986) Adaptation and color matching. *Vision Research*, 26, 1925–38.
Zaidi, Q., Shapiro, A., and Hood, D. (1992) The effect of adaptation on the differential sensitivity of the S-cone color system. *Vision Research*, 32, 1297–318.