Chapter 5

Psychophysical Scaling of Color Attributes: Stimulus and Perception

In Chapter 4 color scaling was considered from a purely psychological point of view. Once attributes were selected, judgments were made using samples of colored materials to exemplify scales of the attributes in conformance with the geometrical form selected for the color solid. In a similar manner it was also possible to study thresholds and suprathreshold color differences without resort to physical measurements. It was of interest to the early psychophysicists to find the relationship between stimulus magnitude and response. It has been indicated before that such relationships can only be valid for tightly circumscribed conditions of viewing and that they can be significantly different for other conditions. Within these limits it is desirable to characterize both stimulus and resulting perception with a high level of accuracy and repeatability. Since the characterization of the stimulus involves physical measurements, progress was linked to progress in photometry and spectrophotometry.

5.1 REQUIREMENTS FOR A UNIFORM PSYCHOPHYSICAL COLOR SPACE

A formula representing a uniform psychophysical color space for trichromats must transform spectral power distributions or reflectances in such a way that

Color Space and Its Divisions: Color Order from Antiquity to the Present, by Rolf G. Kuehni
ISBN 0-471-32670-4

color scales in three dimensions in good agreement with psychological scales are created. Not only must the colors, for example of a color circle, be placed in the correct ordinal order, but they must also be correctly ordered in regard to saturation or chroma, brightness or lightness, and the unit differences in hue, chroma and lightness, must be the same. As we will see, simple psychophysical spaces such as a form of cone response space, the CIE X, Y, Z space, or linearly related opponent color spaces place colors in proper ordinal sequence; however, distances in these spaces are not proportional to psychological distances. Since the beginning of the twentieth century a search has been on to find the transformation that might accomplish this task. It has become evident that a single formula cannot accomplish it for any kind of color difference, or only perhaps with several adjustable variables. We begin by separately considering the three conventional key attributes of psychological object color order: lightness, hue, and chroma.

5.2 POSTULATED RELATIONSHIP BETWEEN PSYCHOLOGICAL AND PHYSICAL MAGNITUDES

Historically postulated relationships between psychological and physical magnitudes are fraught with difficulties. These are due to the absence of a validated mechanism connecting the psychological experiences of color with the physical measurement of spectral power. The problems begin with the definition of the brightness function. Additive results (i.e., where two stimuli mixed together are seen as resulting in a brightness perception that is the sum of the perceptions of the two component lights) are only obtained if the spectral brightness function is determined according to a particular methodology (see below). Brightness and lightness perceptions thus are complex and dependent on experimental conditions.

The CIE colorimetric system, developed in the late 1920s, is based on experiments establishing what lights are matching. It does not have an explicit connection to the appearance of stimuli. Such a connection was usually assumed, however (e.g., see Adams, 1942). The relative success of this assumption has strengthened the belief in its applicability. Today models of color vision are usually expressed in terms of cone sensitivity functions. These are believed to be linearly related to color-matching functions, and vice versa. Models of color vision have become more sophisticated by considering surrounds and other experimental conditions, but the fundamental assumption of the direct relationship between color matching and color appearance remains.

Brightness and lightness and their differences are expressed in terms of luminance or luminous reflectance, even though this does not consider the spectral brightness component of the Helmholtz-Kohlrausch effect (see below). Only the OSA-UCS system has included this effect. Conventionally hue and chroma are believed to be functions of properly scaled opponent signals derived from CIE tristimulus values, hue the ratio and chroma the euclidean sum (see Section 5.6). These are empirical approximations with a

degree of validation from the correlation with visual data, as we will see. To bring such models into close agreement with psychological results requires extensive fine-tuning and application of auxiliary functions, as Chapter 6 will show. The following sections discuss the development of these implied relationships in some detail.

5.3 PHOTOMETRY AND BRIGHTNESS/LIGHTNESS

For the visual sense the earliest quantitative studies comparing stimulus to response involved brightness. Photometry, the quantitative measurement of light, has been of interest since the Renaissance. In its more modern form it was developed by Bouguer and advanced by Lambert and Benjamin Thompson, the Count of Rumford (1753–1814). The initial equipment consisted of simple wedges that allowed visual comparison of two different light sources under controlled conditions. A curious, very early illustration of what appears to be a photometric measurement[1] is that by the celebrated painter Rubens used as the frontispiece for Book V of d'Aguilon's text of 1613 *Opticorum Libri Sex* (Fig. 5-1). Here the effect of distance of light source on the perceived brightness of the patches on the screen is being studied.

Bouguer made quantitative measurements by comparing a test light source against a candle as the reference source. The first step toward making measurements more reliable was to use an artificial reference light source other than a candle. This resulted, beginning in the late eighteenth century, in the

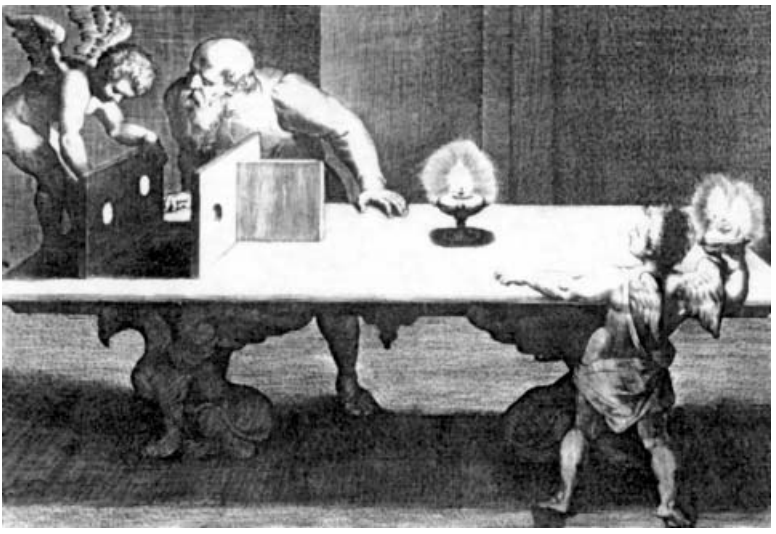

Fig. 5-1 *Rubens's, frontispiece illustration to Book V of d'Aguilon's* Opticorum libri sex *depicting an experiment in measuring light intensity as a function of distance of the lights, 1613.*

development of oil lamps, pentane lamps, and the Hefner lamp, burning various hydrocarbons (used until the 1940s).[2] The human observer began to be replaced by photoelectric cells beginning in the early twentieth century. Lambert had already envisaged this development by describing in *Photometria* in concept a piece of equipment that could measure light analogous to temperature measurement with a thermometer.[3] For such measurements to be in agreement with visual measurements the correspondence between the spectral sensitivity of photoelectric cells and that of the average human had to be established.

The discovery of rods and cones of the human visual system in 1828 by the German physiologist G. R. Treviranus (1776–1837) and the discovery of visual purple in 1877 by F. Boll resulted in extended efforts to determine their purpose and activity.[4] The duplicity theory of the double function of the retina, according to which the cones mediate daylight vision and the rods night vision, was proposed by the German anatomist M. J. S. Schultze (1825–1874) in 1866 and expanded by Helmholtz's student Kries. Measurements of the luminosity of spectral colors at different levels of illumination were made in Helmholtz's laboratory by König and C. Dieterici (1884, 1892). König was able to show the close agreement among an absorption curve he had measured for visual purple, the luminosity curve of a person with monochromatic vision established under daylight conditions, and the luminosity curve for an observer with normal vision determined at very low light levels. This was a clear indication of the existence of the postulated second visual system based on rods and operational for night vision in observers with normal color vision. König and Dieterici obtained a different spectral sensitivity curve for daylight vision. Comparison of the sensitivity of photoelectric cells and of humans indicated large discrepancies, and in order to bring measurements into agreement with average visual results, filters had to be interposed between light source and the photoelectric cell.

A new type of photometer, the flicker photometer, appeared in the early twentieth century and was intensively investigated by H. E. Ives (1912). The original method of obtaining flicker was to rotate a partial disk in front of a light. Depending on the disk sector and the speed of the disk, more or less flicker is observed. If the occluding portion is half of the disk, at the critical flicker frequency at which the flicker disappears the perceived intensity of light is exactly half the intensity when viewing it unobstructed. By matching the brightness of chromatic flickering fields against a steady field of white light, the luminous efficiency of chromatic lights can be measured. Ives's studies showed that brightness measured in this manner is additive. Additivity is not observed when brightness of lights of different color is measured by comparing steady fields. Given the state of computing at the time, additive functions were very desirable. Important determinations of spectral brightness using flicker photometry were reported in 1923 by K. S. Gibson and E. P. T. Tyndall, working at the U.S. National Bureau of Standards. An average sensitivity curve was recommended by them to the CIE and standardized largely unchanged

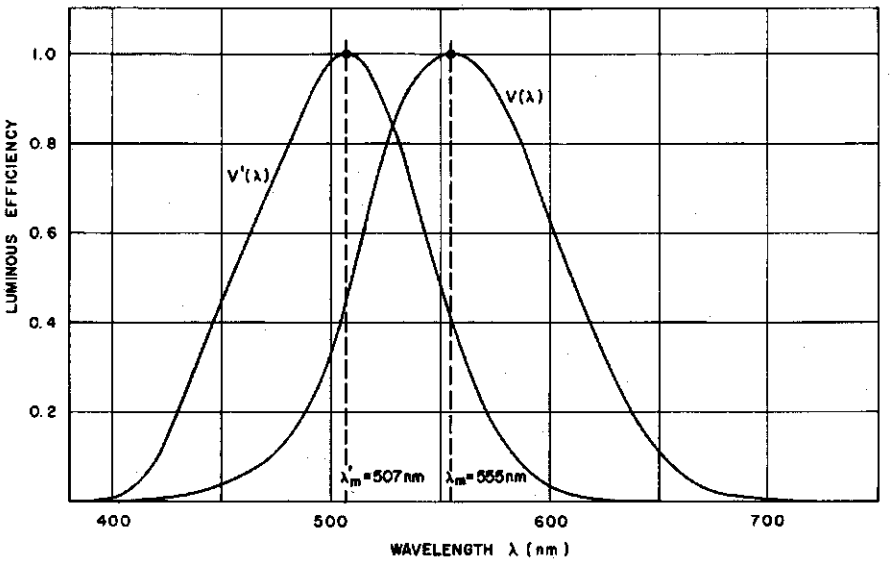

Fig. 5-2 *CIE standard scotopic (low light, V'_λ) and photopic (daylight level, V_λ) luminosity curves. The wavelengths of maximum sensitivity are indicated. From Wyszecki and Stiles (1982).*

by that organization in 1924 as the CIE standard spectral luminous efficiency function $V(\lambda)$ (CIE, 1924), representing the average photopic observer (daylight vision). A comparable scotopic standard observer (low-light vision) was defined in 1951. The spectral sensitivity functions of the two observers are illustrated in Fig. 5-2. The CIE luminosity function has been criticized after its standardization (e.g., see Stockman and Sharpe, 1999), and in 1988 the CIE offered a revised version based on proposals by Judd in 1951 and by J. J. Vos in 1978 (CIE, 1990). The 1924 function nevertheless is an integral part of the CIE colorimetric system and continues to be widely used.

5.4 THE COLORIMETRIC SYSTEM

The idea of trichromatic color vision was originally put forth by Palmer in 1777 and revived by Young in 1802. Important support was later provided by the work of Helmholtz, Grassmann, and Maxwell, as discussed earlier. In the 1850s Helmholtz estimated the spectral sensitivity of the human color vision system in agreement with Young's theory (Fig. 5-3; Helmholtz, 1860). In 1860 Maxwell provided the first set of measured functions derived using a visual colorimeter (Fig. 5-4). Detailed measurements were made by König and Dieterici in 1892, using an advanced König-Helmholtz spectral colorimeter (Fig. 5-5). Similar measurements were made also by Abney and reported in 1914.

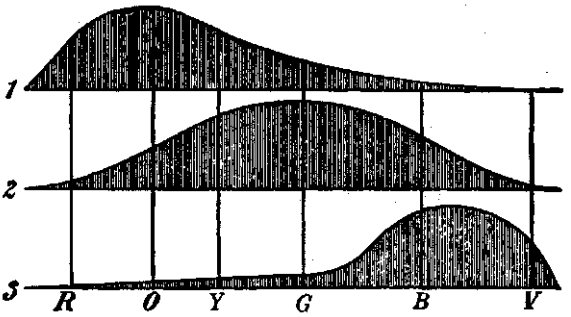

Fig. 5-3 *Helmholtz's sketch of estimated spectral sensitivity of three fundamental color vision processes, 1860. The letters refer to major hues of the spectrum, from red (*R*) on the left to violet on the right.*

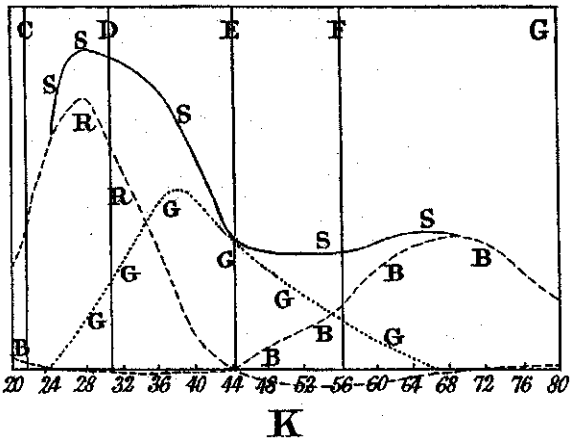

Fig. 5-4 *Spectral curves of the three fundamental processes of color vision* R, G, *and* B, *and the sum curve* S *as determined by Maxwell, observer K. The wavelength scale is in arbitrary units;* C–G *identify Fraunhofer lines in the spectrum, 1860.*

Between 1920 and 1930 J. Guild (1930) and W. D. Wright (1928–29) in England independently built improved equipment to measure these functions, and the color-matching functions of various observers were determined. These data were considered by the CIE for standardization. Following a suggestion by Judd, linear transformations were calculated so that two of the implied primaries were located on a line with zero luminance (the alychne). This resulted in the third primary having a spectral function identical to the CIE luminance function. In this form the functions were standardized as the CIE 1931 2° standard observer color-matching functions (applicable to a field of view of 2°, Fig. 5-6; CIE, 1931).[5] In 1964 the CIE 1964 10° standard observer color-matching functions were added, applying to a field of view of 10° (CIE, 1964) and based

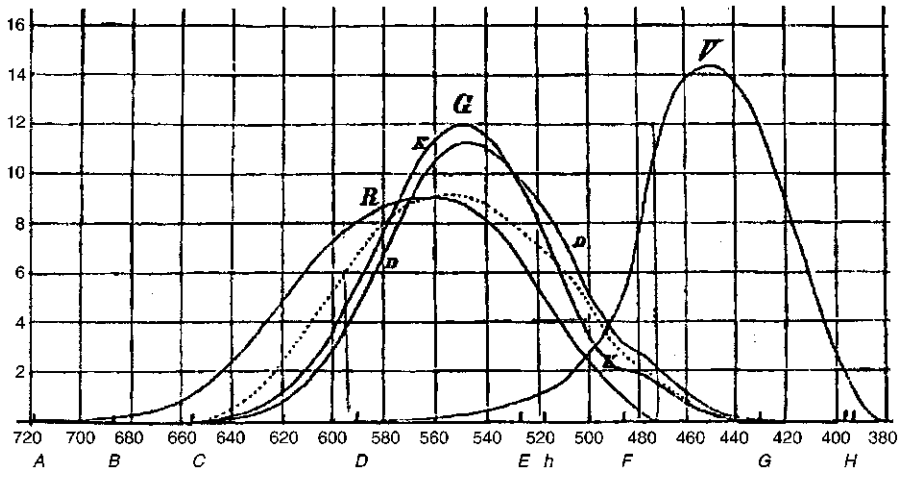

Fig. 5-5 *Measurements of their own fundamental color vision processes* R, G, *and* V *(violet) by König and Dieterici. The dashed line represents the* G *function of an observer with anomalous color vision, 1886. Note that the wavelength scale is in reverse. Letters along the wavelength scale denote major Fraunhofer lines.*

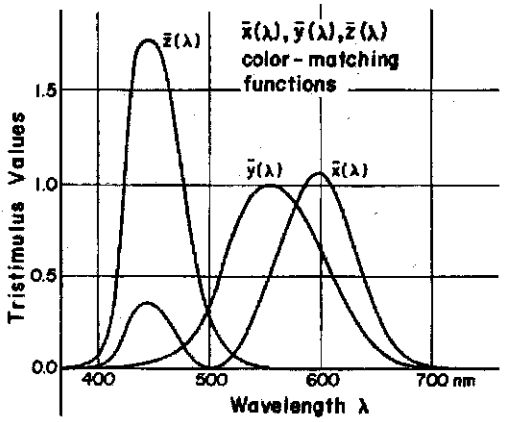

Fig. 5-6 *CIE spectral color matching functions of the 2° standard observer. They have been linearly transformed from the measured functions. From Wyszecki and Stiles (1982).*

on measurements by W. S. Stiles and J. M. Burch (1955) and by N. I. Speranskaya (1959). The two sets of functions differ somewhat.[6] With these definitions of standard observers, as well as definitions of illuminants and of the reflectance factor defining the reflecting properties of object colors, a system was available for quantitative colorimetry. In this system any colored material as viewed under a standard light source by a standard observer is expressed with three figures identified as the tristimulus values X, Y, and Z, where Y is

identical to the luminous reflectance for object colors (luminance for lights). The X, Y, Z space is a nonuniform psychophysical color space, taken to be euclidean, relating measured reflectance of a sample viewed in a standard light to a specific point in the space. Tristimulus values in case of object colors are normalized sums of the reflectance function weighted by the spectral power distribution of the light source and the three color matching functions $\bar{x}$, $\bar{y}$, $\bar{z}$. With one dimension in this space representing (flicker) luminance, the other two must represent in some fashion the chromatic components from which hue and chroma are derived. In the CIE colorimetric system the psychophysical definition of hue is the dominant, or complementary wavelength, of saturation colorimetric purity as expressed in the CIE chromaticity diagram, a particular version of chromatic plane defined by the chromaticity coordinates x and y. Their definition is as follows:

$$x = \frac{X}{X+Y+Z} \qquad \text{and} \qquad y = \frac{Y}{X+Y+Z}\,. \tag{5-1}$$

In the chromaticity diagram spectral colors fall on a horseshoe-shaped function, with saturated nonspectral purple colors falling on a straight line connecting the ends of the spectral function (Fig. 5-7). All possible object and light colors fall on or within the outline of the diagram. The chromaticity diagram is also not visually uniform. A color space can be constructed with the axes x, y, Y. It is commonly used to represent color relationships. An example is the Rösch-MacAdam solid illustrating on its surface optimal object colors for a given standard observer/illuminant combination (see Fig. 2-46). The X, Y, Z space and the chromaticity diagram allow the determination of quantitative relationships between colors in a linear system, but without consideration of surrounds. The locus of spectral colors in X, Y, Z space is illustrated in Fig. 5-8.

The CIE colorimetric system has not escaped criticism and has been even ridiculed.[7] Despite its weaknesses it remains a central aspect of color science and its technical applications. Among the reasons are its noticeable technical successes, the lack of better proposals, and resistance against change. True progress, if possible, requires knowledge of the contents of the black box connecting physical stimuli with the experience of color.

5.5 CONE RESPONSE SPACE

Physiology has identified three cone types in the human retina, in agreement with the postulates of trichromatic vision (but recall the comment about females with the genetic potential for four cone types in Chapter 1). These three cone types are presumably the only transducers of information responsible for color vision (rods may contribute to color vision under certain limited circumstances). Along the visual path in the brain many transformations of

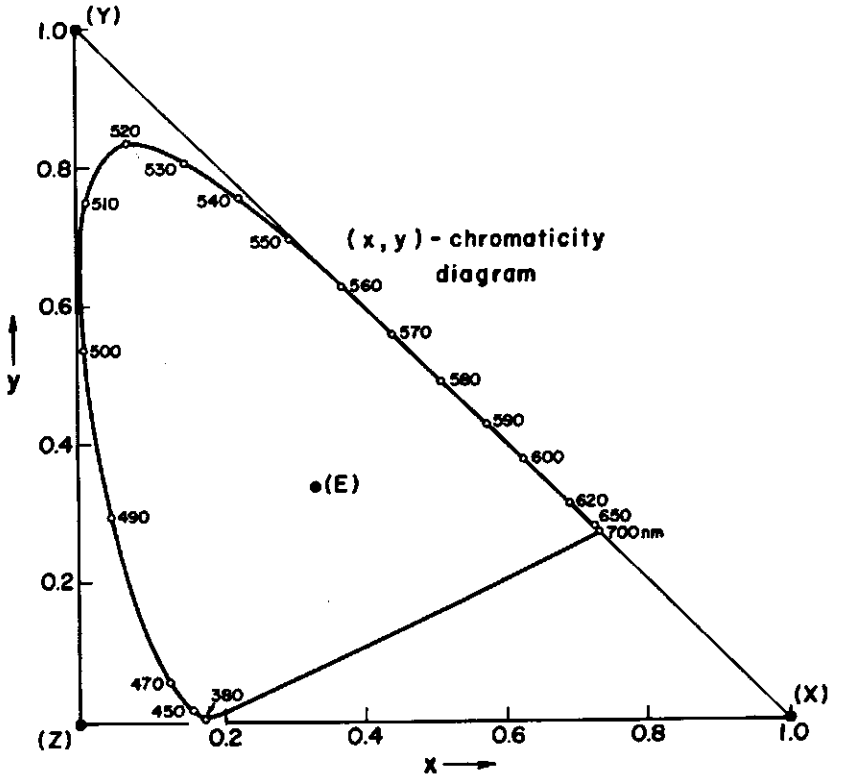

Fig. 5-7 *CIE chromaticity diagram with spectral trace and the nonspectral purple colors for the 2° standard observer.* E *represents the equal energy illuminant. From Wyszecki and Stiles (1982).*

the information generated in the retina take place, all derived from cone responses. Cone sensitivity and cone response to specific light energy are therefore important pieces of information in an attempt to link physical stimulus and psychological response. Attempts to measure spectral absorption functions of cone pigments began in the middle of the twentieth century, with refinement continuing into the present. However, light radiation undergoes changes in the eye before being absorbed by the cones of the retina, changes that need to be accounted for. They are due to the media in the eye from the outer surface to the cones. For a comparison between stimulus and perception, cone response functions taking into account intraocular absorptions are therefore essential. In vivo measurements of human cone response are not practical, and the data currently accepted are based on absorption measurements of excised eyes, on measurements made on primates, and on the assumption that cone response functions are linearly related to color-matching functions. Proposals for standard cone sensitivity functions have been made

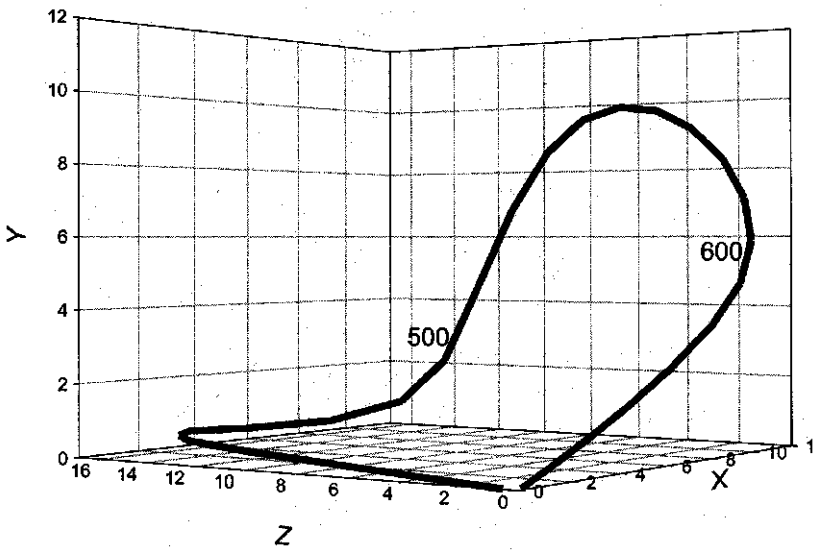

Fig. 5-8 *Locus of spectral colors in the CIE* X, Y, Z *space, beginning at 380 nm and ending at 750 nm (wavelengths 500 and 600 nm are identified).*

by V. C. Smith and J. Pokorny (1975), by Vos (1978), by A. Stockman et al. (1993), and by Stockman and L. T. Sharpe (1998). The latest proposals are based on the Stiles-Burch color-matching functions. There are relatively small differences between the three sets of data that for the current purposes do not seem to be of significance, however, even though, as discussed below, they could have a noticeable effect on correlation with hue difference perception. Cone response functions are reported in different ways: linear or logarithmic scales, function peaks at unity, or identical area under the curves. Figure 5-9*a* and *b* illustrate the Smith-Pokorny functions (as derived from the CIE 2° observer data) in linear scale as well as in logarithmic scale. In the past the three functions were often called *R*, *G*, and *B*, for red, green, and blue. In recent years, as more details about the opponent color system became known, it has been realized that this is not an appropriate way of identifying them. They are now called *L*, *M*, and *S* (for long, medium, and short waveband sensitivity). The three functions can be thought of as constituting the axes of a (nonuniform) psychophysical color space. Figure 5-10 illustrates the spectral trace in the orthogonal *L*, *M*, *S* space defined by the Smith and Pokorny functions. None of the axes of this space has a direct relationship to a color attribute. An unresolved issue is the effect of field size on the cone sensitivity functions. If, as said above, cone sensitivity functions are linearly related to color-matching functions, the cone sensitivity functions must be different for the 2° and 10° standard observers. There are reasons why this could be so. Some vision scientists believe that the same transformation

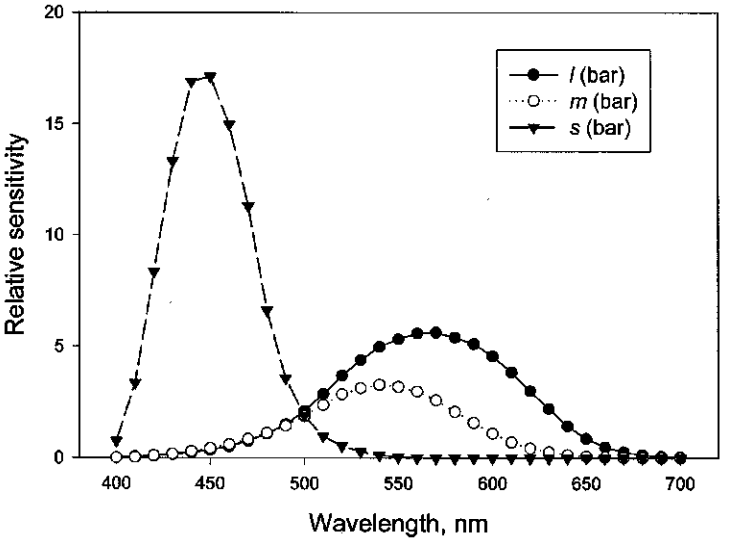

Fig. 5-9a *Spectral cone sensitivity functions according to Smith and Pokorny, calculated for the CIE 2° standard observer, with identical area under the curves, linear sensitivity scale.*

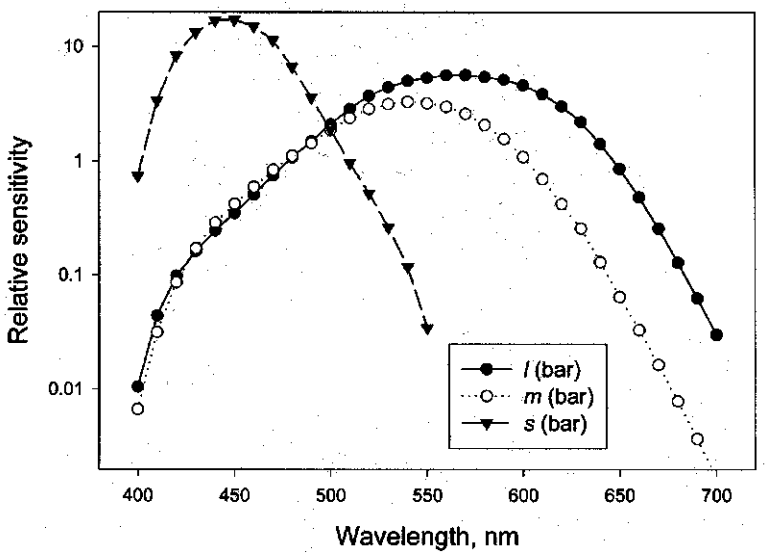

Fig. 5-9b *Spectral cone sensitivities from Fig. 5-9*a *in logarithmic sensitivity scale.*

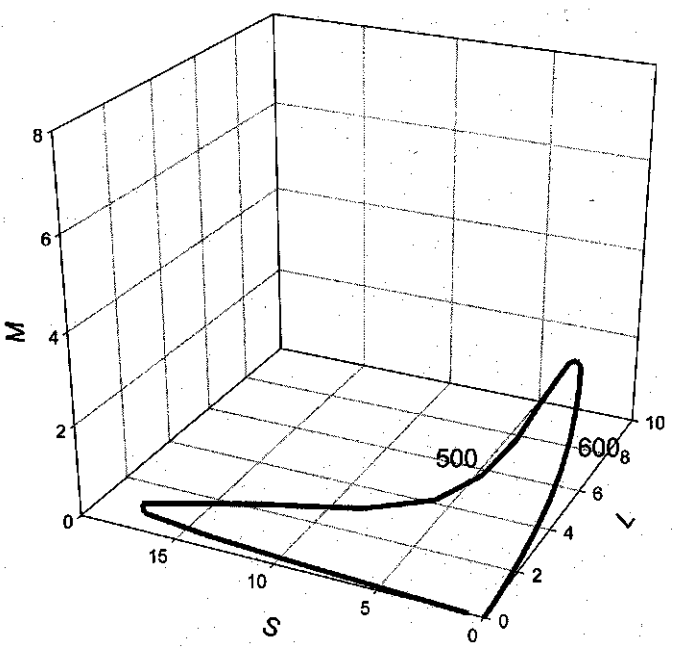

Fig. 5-10 *Spectral trace in the L, M, S space based on the cone sensitivity functions of Fig. 5-9a. The wavelengths 500 and 600 nm are identified.*

formulas are applicable for the 2 and 10 standard observers (Smith, 1999).

Cone responses are essentially in linear relationship with the absorbed light energy up to a luminance of approximately 1000 cd/m^2 (a cloudy sky at noon). Above that, for example, in bright sunlight, saturation effects occur resulting in relatively reduced response.

Individual Variation in Cone Functions and Smoothness of Functions

In recent years a significant effort is under way to develop a new set of cone functions taking into account the latest information on genetic variability of cone expression and using modern technological means (Stockman and Sharpe, 1998, 1999).[8] Genetic research has indicated that the *L* and *M* cones can be expressed in different ways and that some people have a hybrid *L/M* cone (Sharpe et al., 1999). Wavelengths of peak response for *L* cones vary in individuals from approximately 554 to 563 nm. There is less variation in *M* cone expression, with peak wavelengths varying from approximately 532 to 538 nm, and only one normal expression of the *S* cone. There are differences between the sexes in the prevalence of genetic expressions of cones. As mentioned in Chapter 1, it is known that up to 50% of females have the genetic potential

for four cone types. So far no individual actually having four cone types has been identified. All these issues touch on the question of a statistically meaningful average observer and how different sets of data may be influenced by the genetic variability in its observer pool. Such observer pools are likely to also differ in the degree of ocular yellowing and distribution of macular pigment, both affecting in particular the *S* cone sensitivity function. Another issue in connection with cone functions is the question of their intrinsic smoothness. CIE color-matching functions are smooth, and cone functions calculated from these by linear transformation are also smooth. Directly measured individual cone sensitivity functions differ somewhat in shape and appear to have in some cases more or less pronounced "dents" (see Fig. 5-5 for an example). Some of these appear to be due to the degree of macular absorption in individuals, as already König surmised.

5.6 OPPONENT COLOR SPACE

Another version of a psychophysical space is the opponent color space. It attempts to provide a psychophysical model for Hering's opponent color hypothesis of color vision and to describe the zone following the cone absorption level in a multiple-zone model. With an opponent color space the conceptual leap is made that there is a connection between functions predicting which color stimuli produce identical color experiences (color-matching functions) and color appearance. Such a connection is not a priori a certainty. It will be shown that such a space can be a rough euclidean approximation of a uniform color space. However, aside from the issue of euclidean nature there are other discrepancies of detail arguing for a considerably more complex picture of reality than a simple opponent color space can provide.

As mentioned before, according to Hering's hypothesis there are three opponent processes that are responsible for our color experiences: a greenness-redness, a yellowness-blueness, and a whiteness-blackness process. Hering had developed physiological ideas of how such a process could be implemented. He engaged in an epic struggle with Helmholtz pitting the opponent color theory against what is known as the Young-Helmholtz trichromatic theory. After conceptual proposals by Donders and Kries, Helmholtz showed in 1896 how the two zone components could be expressed in mathematically equivalent form, but voiced doubts about the need for a second zone step. The psychologist and Hering supporter G. E. Müller developed a complex three-stage theory in 1896 to be championed in the 1960s by Judd. Schrödinger (1887–1961) showed in 1926 again, using the König and Dieterici spectral sensitivity data, how the trichromatic and the opponent color theory can be combined. From the sensitivity data he calculated spectral opponent color response functions (Fig. 5-11). In 1927 Luther in Germany and, independently, in 1929 Nyberg in Russia proposed what came to be known as the Luther-Nyberg color solid (see Section 2.22, Fig. 2-47). It is based on a (nonuni-

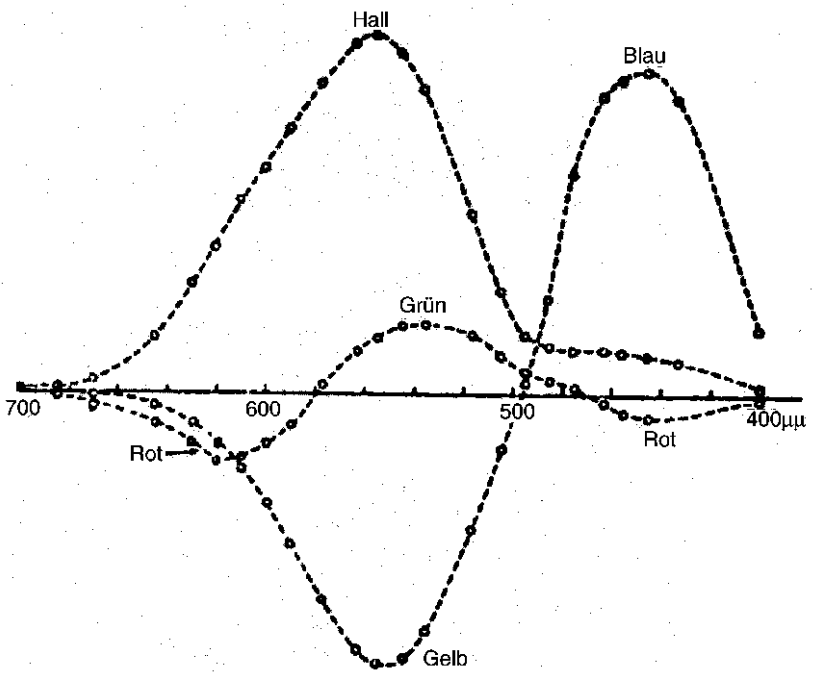

Fig. 5-11 *Schrödinger's opponent color and brightness functions calculated from König and Dieterici's fundamental functions (Fig. 5-5). Note that the wavelength is reversed from the customary presentation, 1926.* Hell *refers to the brightness function. The other two functions represent yellowness-blueness (*Gelb, Blau*) and greenness-redness (*Grün, Rot*).*

form) linear opponent color space with two color moments and a color weight.

In 1923 the American physicist E. Q. Adams proposed a zone theory in which the output from three cone types is modulated and subtracted to result in opponent color signals. Adams expanded on his proposal in 1942 where he showed, using CIE tristimulus data, that a theory such as he had proposed could offer a good model for the Munsell system (Fig. 5-12). There he specifically proposed that the output of the three cone types should be considered identical to the CIE tristimulus values and the modulation of the output identical to that of the Munsell value function. He constructed a "chromance" diagram by subtracting modulated Y values from modulated X and Z values. In 1944 Nickerson and K. Stultz offered a color space and color difference formula based on Adams's proposals which became influential in technology (more details in Chapter 6).

At the same time support for an opponent color step in color vision began to develop on a different front. Around 1940 physiologists began to measure signals by inserting microelectrodes into retinal cells. The Swedish physiologist R. Granit measured spectral response curves in ganglion cells in the retina

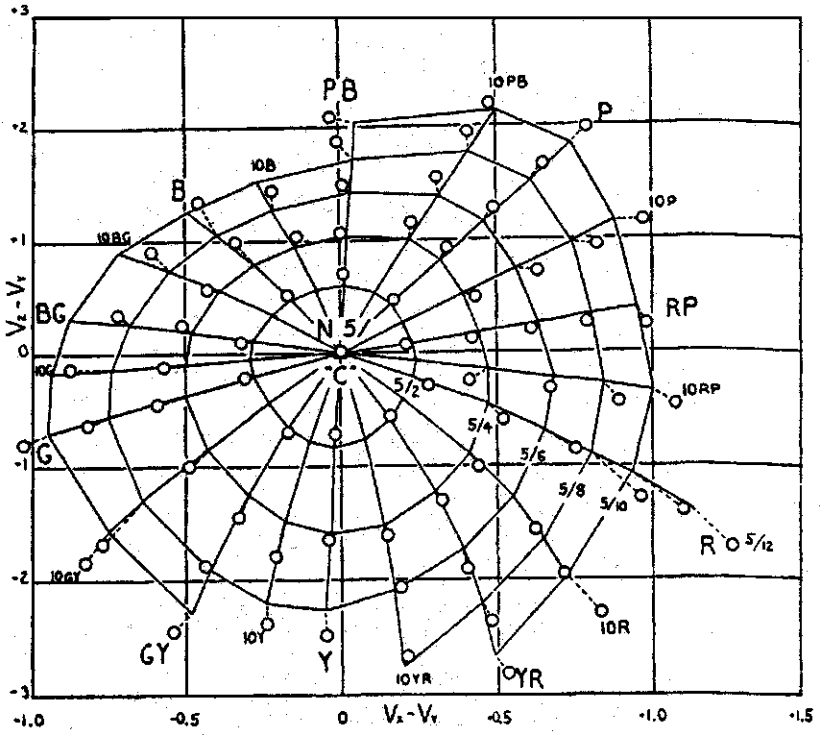

Fig. 5-12 *Munsell colors at value 5 of the 1929* Book of Color *in Adams's 1942 opponent color chromatic diagram. Lines represent the smoothing proposal by Newhall (1940).*

of frogs. His student G. Svaetichin found chromatic opponent cells in fish retinas in 1956.[9] In the 1960s American researchers made measurements from cells of the lateral geniculate nucleus of monkeys and found various kinds of cells with opponent color sensitivity (De Valois, 1965; Wiesel and Hubel, 1966).

On the psychological front the American psychologists D. Jameson and L. Hurvich began in the mid-1950s to use hue cancellation as an experimental method and deduced opponent chromatic response functions from the results. In such experiments, for example, a "blue" stimulus is added to a yellowish color until the perceived yellowness is canceled or neutralized. They were able to show that their personal experimental functions could be reasonably approximated using CIE standard observer color-matching functions (Fig. 5-13) in a manner employed by Adams and by Judd in his interpretation of the Müller theory (Hurvich and Jameson, 1955; see Chapter 6 for more details). In 1964 Hurvich and Jameson also translated into English Hering's posthumously published *Grundzüge der Lehre vom Lichtsinn* (Outlines of a theory of the light sense; Hering 1905–1911).

If hue changes depend directly on the chromatic responses shown in Fig. 5-13, it is apparent that they vary considerably as a function of wavelength.

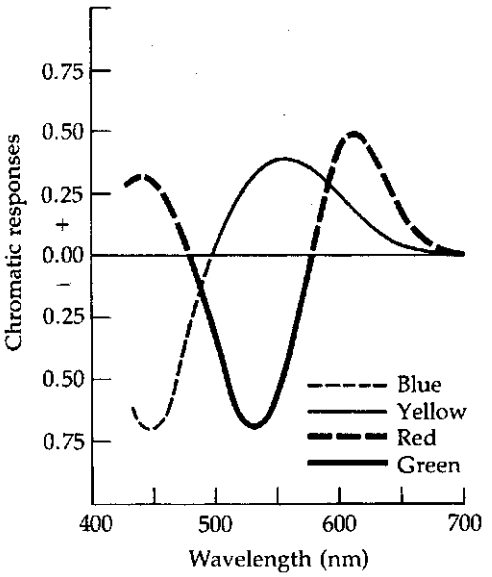

Fig. 5-13 *Opponent color functions of the CIE 2° standard observer calculated by Jameson and Hurvich (1955).*

Near 475 nm the greenness-redness signal is at zero, indicating the presumed location of unique blue for the CIE standard observer. The blueness signal on both sides of that point changes rapidly, as does the greenness-redness signal, indicating rapid hue changes per nanometer increment. This is indeed the case as wavelength discrimination experiments show. At approximately 500 nm is the neutral point of the blueness-yellowness signal (unique green, but see Section 5.10). Also here both signals change relatively rapidly indicating continued high sensitivity to wavelength changes. By this conjecture the complete change from unique blue to unique green takes place within about 25 nm. At about 578 nm the greenness-redness signal passes through its second neutral point resulting in the display of unique yellow. While the yellowness signal changes relatively slowly in this region, the greenness-redness signal changes rapidly on both sides of the neutral point indicating rapid changes toward greenish yellow and reddish yellow away from the neutral point. This is again supported by a high rate of change in wavelength discrimination experiments. In the regions of approximately 440, 530, and 610 nm and above the rate of change is significantly lower, resulting in reduced perceived wavelength discrimination. While a reasonable agreement exists between the form of opponent functions and experimental wavelength discrimination, there is no solid agreement between rate of change data behind these functions and experimental hue difference data, as will be shown below. There is in addition the matter of the wide ranges of wavelengths of the blue, green, and yellow unique hues as determined by individual observers.

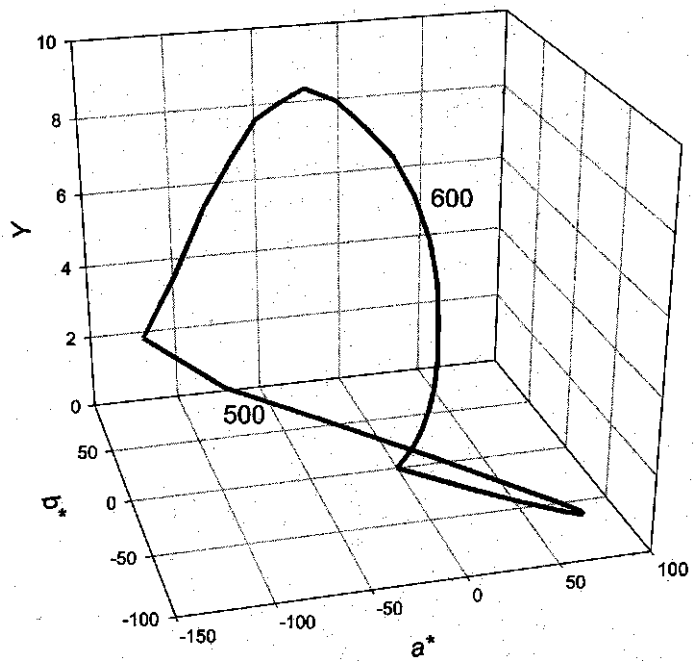

Fig. 5-14 *Spectral trace in the* Y, a*, b* *space, 10° observer. The wavelengths of 500 and 600 nm are identified.*

As Chapter 6 will show, most color vision models developed since the mid-nineteenth century are zone models and a simple version of a zone model color space was recommended in 1976 by the CIE as the CIELAB formula (CIE 1976; see Chapter 6). In this space, considered euclidean and with a cartesian chromatic diagram, colors are identified by chromatic coordinates a^* and b^* and a lightness coordinate L^*. There is also a polar coordinate interpretation where chroma is defined as the radial distance from the origin and hue by the hue angle, thus bringing the system conceptually in agreement with a psychological system of the Munsell type. Experimental work before and since then has indicated that there are surround effects requiring consideration and that CIELAB space is at best a rough approximation of a large difference uniform color space. Figure 5-14 illustrates the spectral colors in the Y, a^*, b^* space.

An important matter usually not considered in models until very recently is that of the interaction of cone responses between test field(s) and surround. Four types of opponent cells involving L and M cone output comparison have been identified (Wässle et al., 1994). All four involve comparison of output of cones from the center of the receptive field of the opponent cell with those from the surround. A similar situation applies to the opponent cells compar-

ing output from *S* cones with that of *L* plus *M* cones. These mechanisms appear to be responsible for test field/surround effects, such as contrast and assimilation, possibly crispening, and others. Correct treatment of these effects is essential for a uniform color space model.

In recent years experimental data have begun to raise doubts about a simple subtractive opponent color system such as proposed by Hurvich and Jameson. New experimental work points to four independent chromatic dimensions and perhaps to a multiplicity of hue detection mechanisms (see below).

5.7 HOW ARE THE *L, M, S* AND *X, Y, Z* COLOR SPACES RELATED?

On a mathematical basis, according to Smith and Pokorny (1975), the two spaces are related by a set of linear transformation equations as follows:

$$\begin{aligned} \bar{l} &= (0.15516\bar{x} + 0.54308\bar{y} - 0.03287\bar{z}), \\ \overline{m} &= (-0.15516\bar{x} + 0.45692\bar{y} + 0.03287\bar{z}), \\ \bar{s} &= 1.00000\bar{z}, \end{aligned} \tag{5-2}$$

where $\bar{x}$, $\bar{y}$, $\bar{z}$ are CIE color-matching functions, originally the Judd and Vos modified 2° observer functions. In this transformation $\bar{l} + \overline{m}$ add up to equal $\bar{y}$ (equivalent of the luminosity function). As mentioned, the same equations have been applied to both CIE 2° and 10° observer data. Slightly different equations have been proposed by other authors.

The inverse relationship is as follows:

$$\begin{aligned} \bar{x} &= 2.920\bar{l} - 3.445\overline{m} + 0.208\bar{s}, \\ \bar{y} &= \bar{l} + \overline{m}, \\ \bar{z} &= 1.000\bar{s}. \end{aligned} \tag{5-3}$$

Both the *L*, *M*, *S* space and the CIE tristimulus space are viewed as orthogonal. It is evident that the two spaces cannot be orthogonal in the same reference frame, but a case can be made that they represent different realities and each may be orthogonal by itself. There is, however, the unanswered basic question if the assumption of orthogonality is valid for either space.

Based on their recordings from parvocellular neurons in the LGN of macaques A. M. Derrington, J. Krauskopf, and P. Lennie (DKL) in 1984 proposed a color space (nonuniform) representative of the signal output of different cell types (see Chapter 6). They located opponent cells that report in terms of *L–M* or the reverse and in terms of $(L + M) - S$ or the reverse. When they plotted the corresponding axes in the CIE chromaticity diagram, they

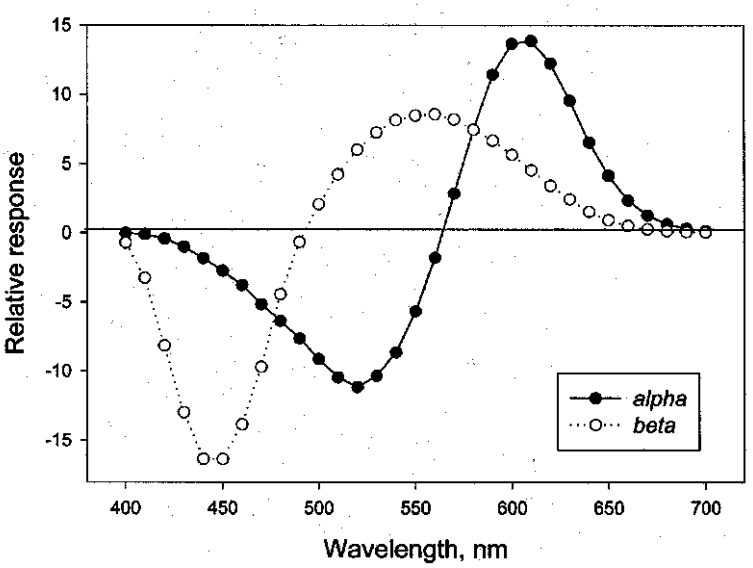

Fig. 5-15 *Balanced opponent functions α and β calculated from CIE color-matching functions. These functions are supported by recordings in the lateral geniculate nuclei of macaques.*

found that the axes do not represent average unique hues. This can be demonstrated by calculating balanced (equal areas under the four curve segments) opponent functions based on cell input reported by DKL:

$$\alpha = 5.673(\bar{l} - 1.98842\overline{m}),$$
$$\beta = \bar{l} + \overline{m} - \bar{s}. \tag{5-4}$$

The resulting two functions α and β are illustrated in Fig. 5-15. The β function is identical to $\bar{y} - \bar{z}$ and therefore has the same form as the Hurvich-Jameson yellowness-blueness opponent color function. The α function is different from the greenness-redness function because it lacks positive values in the short-wave range, considered indicative of the reappearance of redness at short wavelengths. The α function can therefore not be a greenness-redness expressing function (in the sense that green and red are represented by the unique hues), and the DKL space is not a space in alignment with the psychological space (e.g., see Lee 1999). It is important to be aware that the measured functions apply to macaques and that they are from the lateral geniculate nucleus: details in the human visual system may differ and additional transformations in the cortex are known to take place. Recently it has been found that in the primary visual cortex at the back of the brain, most cells receive input from all the different opponent cell types in the lateral geniculate nuclei (De Valois

et al., 1997). On that basis a mechanism for S cone input into the greenness-redness opponent system appears to exist. Equation (5-3) indicates that the definition in terms of cone sensitivities of the $\bar{x}$ color-matching function includes S cone input. It is interesting to note that axis rotation required to provide a result equal to a balanced linear opponent chromatic diagram (based on CIE 1931 tristimulus values) can be achieved by the following equation (Kuehni, 2000a):

$$a = f(\alpha - 0.56\beta),$$
$$b = \beta, \tag{5-5}$$

where a and b are linear opponent color functions, f is a normalization constant with the value 0.989. The factor for β is close to 0.5. It will be shown later that this factor varies considerably for optimal modeling of various sets of visual data. The subtractive step is reminiscent of the second transformation in the G. E. Müller theory (see Chapter 6).

5.8 EXPRESSING PSYCHOLOGICAL SCALES IN PSYCHOPHYSICAL SPACES

Historically the X, Y, Z space (and its predecessors) has been used as a basis to express psychological scales in a psychophysical space. The fact that one of its dimensions is aligned with brightness and the other two therefore with chromaticness seemed a good reason to expect reasonable agreement if the functions expressing $\bar{x}$ and $\bar{z}$ are correctly selected. They were selected so that they had only positive values (the experimentally determined color-matching functions have stretches of negative values). There is, of course, no a priori reason why a color-matching space should be in agreement with, or easily modifiable into, a color appearance space, and as will be seen later, there are important discrepancies.

Lightness Scales Expressed in Terms of Luminous Reflectance

Seemingly the first lightness scale that through reconstruction can be expressed in terms of luminous reflectance is Glisson's gray scale. Figure 2-9 shows the progression of CIELAB L^* lightness values (derived from luminous reflectance) as a function of the grades of the reconstructed scale.

Insufficient quantitative data are available to deduce lightness scales in terms of luminous reflectance for the Mayer, Lambert, Runge, or Chevreul systems. As discussed in Chapter 4, Plateau in 1872 concluded that ratios were applicable, based in part on the apparent preservation of the contrasts in etchings in different lights, and that as a result the relationship would be linearized by a power function. Delboeuf, in an elaborate experiment of halving the

visual distances between two grays independently, concluded that a logarithmic formula provided the best fit to his data. Plateau received confirmation in 1887 from experiments by Breton who found a square root power relationship to be applicable. But, as we have seen, the Weber-Fechner law was the generally accepted paradigm for some 100 years.

Ebbinghaus's eight-grade gray scale had the following relative luminance values: 100, 44.4, 21.1, 10.3, 5.8, 4.4, 2.0, 1.0 (Ebbinghaus, 1902). This series is linearized optimally with a 0.02 power. The resulting scale is not quite uniform and does not quite represent a Fechner type log scale. Ridgeway's gray scale incorporated into his 1912 color atlas had, by definition, a logarithmic step relationship as did Ostwald's gray scale of 1917.

Priest, Gibson, and McNicholas investigated the relationship between Munsell gray scale steps in the 1915 atlas and luminous reflectance values (Priest et al., 1920). They found that the relationship is best described by a square root power function. Taking the 1919 National Bureau of Standards measurements of the neutral samples of the 1915 *Atlas of the Munsell Color System*, as well as the 1926 diffuse measurements of the same samples (Gibson and Nickerson, 1940; see Fig. 7-5), the optimal linearization (minimal coefficient of variation) is found to be based on an 0.42 power. However, the function plotting reflectance against value step is not smooth. It is composed of two segments intersecting between value steps 4 and 5. The lower segment is optimally linearized with a power of 0.25, the upper with one of 0.48. This appears to be an indication of lightness crispening (see below), as one would expect, and points to a surround with a luminous reflectance of approximately $Y = 20$.

In 1922 Adams and his colleague P. W. Cobb studied the effect of surround on perceived brightness and developed a formula (in the version reformulated by Judd) that recognizes the effect of the luminous reflectance of the surround on the results:

$$V = R\left(\frac{Rb + 100}{R + Rb}\right), \tag{5-6}$$

where V is lightness on a scale of 0 to 100; R is luminous reflectance in percentage of the sample, and Rb luminous reflectance of the surround.

A. E. O. Munsell, L. L. Sloan, and I. H. Godlove continued work on the Munsell value scale and described in 1933 results based on a combination of "just noticeable differences" as well as scale-halving experiments. Their scale is similar to the Munsell *Atlas* and 1929 *Book of Color* scales. It is optimally linearized by a 0.34 power but consists of two branches with the lower linearized with an 0.50 power and the upper with an 0.20 power. Godlove (1933) compared it to the Adams-Cobb data and fit the following empirical formula to the Munsell data, a formula described as "intermediate between those of the Fechner-Delboeuf [logarithmic] and Plateau-Munsell [power function] equations":

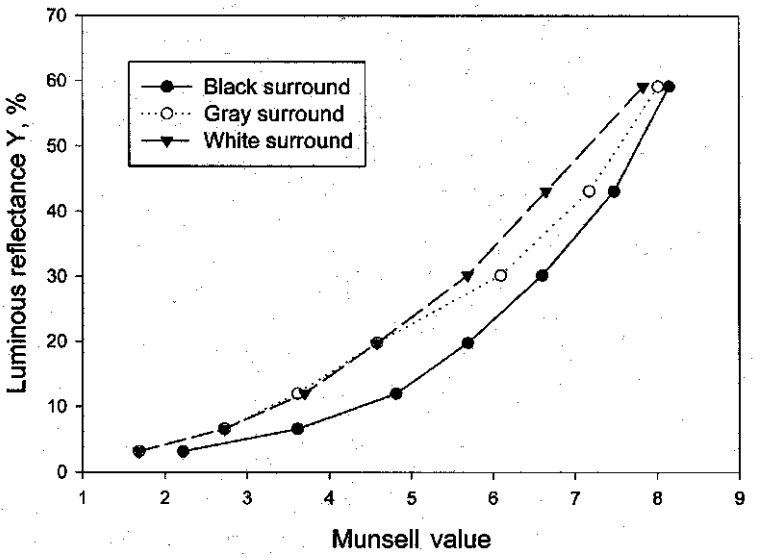

Fig. 5-16 *Results of visual determination of Munsell value of Munsell value scale samples viewed against a white, mid-level gray, and black background. After Newhall et al. (1943).*

$$V^2 = mR + nR^2, \tag{5-7}$$

where V is the Munsell value, m and n are constants, and R is luminous reflectance.

In the studies leading up to the Munsell Renotations the value scale was evaluated against a black, a middle gray, and a white background. The results differed significantly (Fig. 5-16). The optimal power for the results against the black background is cube root against the white background square root. As expected, the results against the gray background show two branches, indicating the lightness crispening effect. The committee working on the Renotations decided to average the results in some fashion, and the final, smoothed curve was mathematically described with a quintic function developed by Judd:

$$Y = 1.2219V - 0.23111V^2 + 0.2395V^3 - 0.021009V^4 + 0.0008404V^5, \tag{5-8}$$

where V is the Munsell value (Newhall et al., 1943). Glasser and colleagues showed in 1958 that this function is closely matched by a cube root function.

Figure 5-17 illustrates lightness scales based on logarithmic (Weber-Fechner) and cube root functions as well as an Adams-Cobb function with a surround luminous reflectance of 35%. It indicates that power functions have less modulating effect than the Weber-Fechner function and that an Adams-Cobb function with a surround luminous reflectance of 35% is roughly comparable to a cube root function.

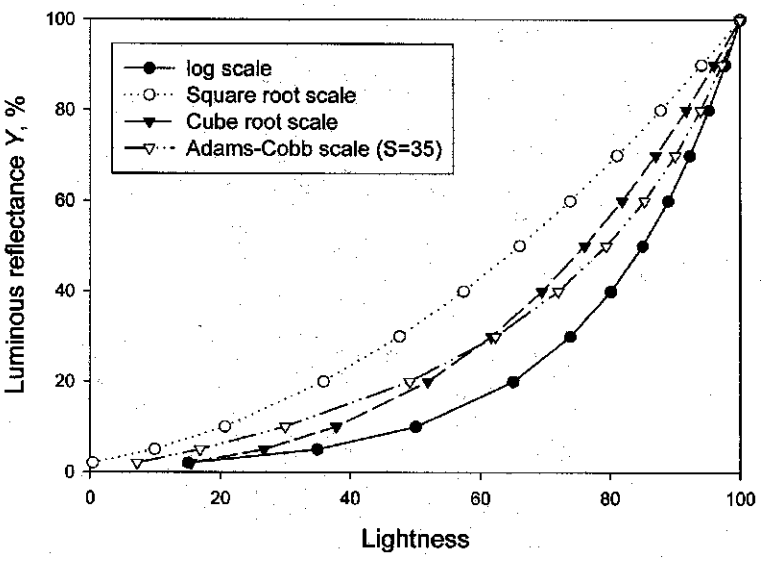

Fig. 5-17 *Luminous reflectance* Y *plotted as a function of lightness according to logarithmic, cube root, square root scaling, as well as the Adams-Cobb lightness scale for a surround of luminous reflectance* Y = *35.*

As mentioned before, in 1933 Schönfelder investigated the effect of surround variation on the precision of color matching. He found that precision is highest if the surround color is intermediate to the standard and the match fields. This applies not only to lightness but also to chromatic differences, and we will encounter what Judd called Schönfelder's law again later. Schönfelder's results were in 1938 confirmed by K. J. W. Craik for achromatic colors and in 1952 by W. R. J. Brown for chromatic colors.

In 1964 Jameson and Hurvich investigated brightness scaling, and they introduced a subtractive constant that was explained as correcting for simultaneous contrast effects:

$$V = bx^p - V_0, \tag{5-9}$$

where V_o is the subtractive constant. In the same year T. Kaneko investigated lightness of achromatic colors in form of color chips as a function of surround lightness. He found, in agreement with Schönfelder, that differences between gray scale grades were perceived largest if the surround lightness falls between the lightnesses of the samples being compared. When the sample lightness is higher or lower than the surround lightness, a larger increment in luminous reflectance is required for a perceived difference of the same magnitude. He also found that the effect depends to some degree on the size of the test fields compared to the surround. With larger test fields the crispening effect was reduced; that is to say, the magnitude of the lightness crispening effect appears

to be a function of the test field size, the effect being strongest for small fields. Kaneko constructed gray scales valid for a given surround lightness (Kaneko, 1964). Adams and Cobb had also noticed the described effect, but they believed it to be limited to threshold level differences. Kaneko found it also to apply for much larger lightness differences. It was H. Takasaki who gave the effect the name *lightness crispening*. Kaneko developed a formula consisting of two simultaneous equations to model his results.

Takasaki conducted in 1966 extensive investigations of the effect of surround lightness on perceived lightness of test fields. He confirmed contrast as well as crispening effects. His complex formula was shown by C. C. Semmelroth in 1970 to be reducible to

$$V = R^m + k\,|R - Rb|^n, \tag{5-10}$$

where V is the perceived lightness, R is the luminous reflectance of the sample, and Rb that of the surround, m and n are exponents, and k is a constant. The equation applies for situations where R is equal or larger than Rb. In the opposite case the +sign is replaced with a –sign. For various experimental data sets $m = 0.4$, $n = 0.2$, and $k = 0.65$ have been found to provide good fits.

In 1975 Judd and Wyszecki constructed an instructive graph of the effect of lightness crispening on perception of value differences in the Munsell value scale, as described by equation (5-10) (Fig. 5-18). Explanatory example: the lightness difference between two Munsell samples $V_s = 2$ and $V_s = 3$ when viewed against a surround of value 7 is approximately $1.65 - 0.7 = 0.95$ units. When viewing the samples against a surround of value 3, the corresponding difference is $3.0 - 1.2 = 1.8$ units, or nearly twice perceived size. More recently lightness crispening has been measured and described by Whittle (1992).

The effect of surround on perceived lightness is demonstrated with the results of an experiment with 22 observers. On a video screen in a dark room, successively against black, white, and gray ($L^* = 45$) backgrounds, the observers adjusted the luminances of eight separated fields so that all steps looked perceptually equidistant. Grade 4 had against the gray surround the same lightness as the surround itself. For the other two surrounds the luminance of grade 4 was adjusted so that it had the same lightness appearance as in the gray surround. The average result in terms of CIELAB L^* lightness is illustrated in Fig. 5-19. The inflection of the crispening effect at the luminance of the surround is clearly visible. It is apparent that the L^* scale is not applicable to any surround lightness. Similar results were obtained with chromatic fields. Lightness crispening has been investigated by Lübbe (1999) and identified in small color difference data by Kuehni (2001b). In agglomerations of data from different experiments made with different surrounds, lightness crispening is obscured. In several individual data sets it is clearly detectable. It was, for a single surround, measured by Chou and colleagues (2001). For small color difference formulas, several different weightings of the CIELAB lightness scale have been proposed since the mid-1970s, as shown in Chapter

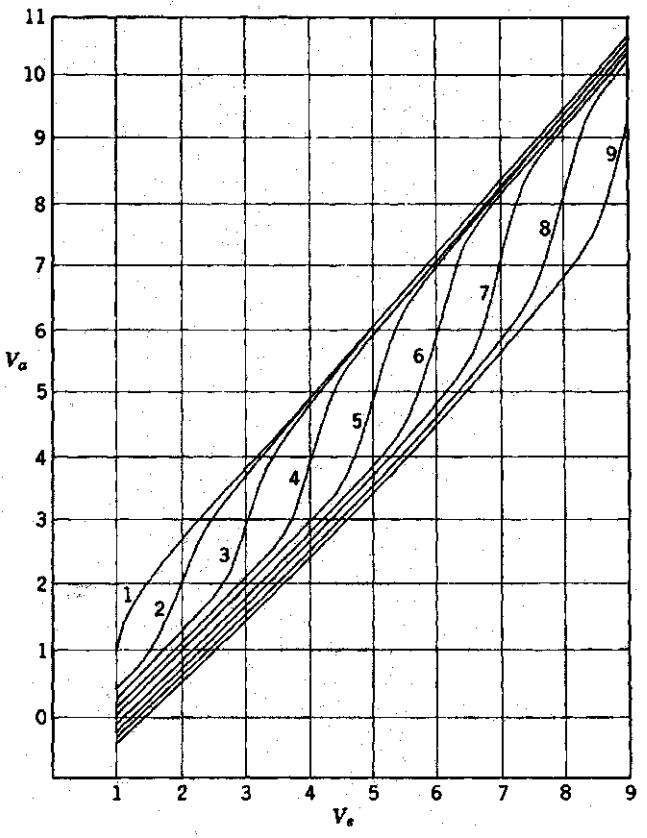

Fig. 5-18 *Graph developed by Judd and Wyszecki from a table by Semmelroth (1971) depicting the effect of lightness crispening on perception of value differences in the Munsell Renotation value scale. From Judd and Wyszecki (1975).* V_s *is the nominal Munsell value,* V_a *is the background adjusted Munsell value.*

6. Surprisingly few formulas consider surround lightness explicitly (Lübbe, 1999; Kuehni, 2001d) where it has improved the correlation between visual and calculated color differences (Kuehni, 2001b).

Some researchers recently began to use the term *masking* for the crispening effect. Stimuli producing a strong contrast to the surround are said to produce masking and require larger increments for detection than those at lower contrast (Eskew et al., 1999). Aside from luminance or luminous reflectance masking there is also chromatic masking, discussed below. In such experiments masking is reported to produce threshold elevations that can be approximated with a power law with exponents from 0.5 to 1.0. Such masking experiments typically involve only brief exposures to the stimuli and other controls that minimize adaptation and other effects. The results therefore cannot be expected to closely match those of experiments involving object color samples with unrestricted exposure.

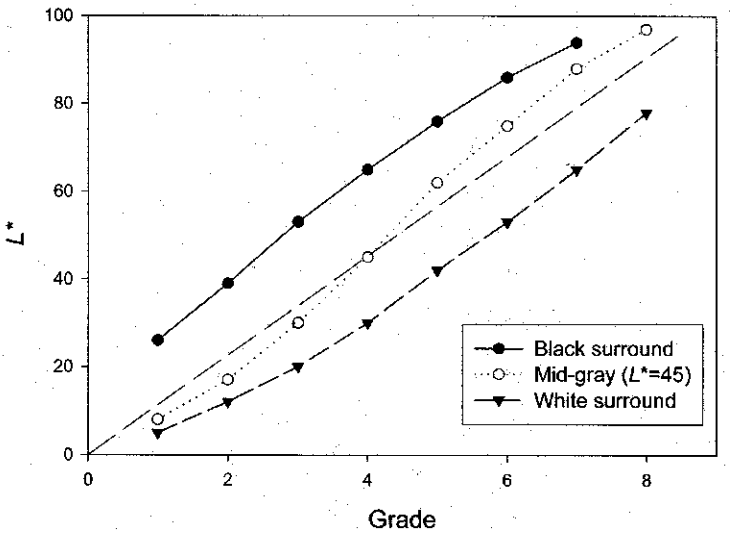

Fig. 5-19 *Average settings expressed in* L* *values of 22 observers of the luminance of eight grades of achromatic colors against three surrounds so that steps between the grades appear equidistant.*

The CIELAB formula of 1976 enshrined the Munsell value function in a cube root format to connect luminous reflectance with perceived lightness. The OSA-UCS formula, as will be seen in Chapter 6, used the Munsell value function as a basis but with adjustments that consider lightness crispening and the Helmholtz-Kohlrausch effect, the latter to be discussed presently.

The gray scale of the Swedish Natural Color System was obtained by assessment of a series of gray samples and their percentages of innate ideas of the observers of ideal white and black. The Adams-Cobb formula with a surround luminous reflectance of $Y = 56$ was found to fit the experimental data well (Hård et al., 1996).

It should be mentioned here that a parallel process, chromatic crispening, exists. According to this effect the smallest stimulus increment for a criterion chromatic difference is required if the chromaticities of the two test samples straddles the chromaticity of the surround (see Chapter 8).

The work of Gilchrist and colleagues (1999) and of Purves and colleagues (see Purves and Lotto, 2002) in regard to the complex relationship between luminance and brightness/lightness was mentioned in Chapter 4. In the Munsell and similar scales the luminance ratio between white and black is close to 100:1, and according to the Gilchrist group's findings, luminance scale compression takes place under these conditions, just as it is found in most data of that kind.

Heterochromatic Brightness/Lightness

When luminance measurements of colored stimuli began to be made routinely, it became evident that the perceptual brightness of constant luminance stimuli was higher than that of an achromatic color with equal luminance. The chromatic stimulus appears to glow compared to the achromatic one. Expressed conversely for object colors, highly chromatic objects having the same perceived lightness as a gray object have a reduced luminous reflectance. Colors of different hues have a different degree of color glow. The result of these findings is that modulated luminous reflectance is not a good predictor of perceived lightness of chromatic stimuli. Helmholtz remarked upon this effect (1860), and it was investigated by V. A. Kohlrausch in 1935, among others, and has become known as the Helmholtz-Kohlrausch effect (HKE). The implication is that an effect in addition to that described by the luminous efficiency function contributes to perceived lightness. The effect is often quantitatively described by the B/L ratio, where B is the luminance or luminous reflectance of the reference stimulus, usually achromatic, and L that of the test stimulus (e.g., see Wyszecki and Stiles, 1982). It is of considerable importance for the photometry of colored lights but also for the definition of a uniform color space.

Interestingly and surprisingly, in the Munsell 1915 atlas the majority of the most highly chromatic chips have a luminous reflectance that is higher than that of the gray with the same visually judged Munsell value. In the Renotations value was defined in terms of luminous reflectance and all colors of a given value have the same luminous reflectance value. The Re-renotations incorporate the HKE and colors of the same value differ in luminous reflectance. The same applies to the OSA-UCS system where the HKE also has been incorporated. In the 1950s and 1960s important experimental studies of the effect were made. Among these is the Wyszecki-Sanders work of 1964 with 20 observers using a visual colorimeter, and an experiment by the Committee on Uniform Color Scales of the Optical Society reported by Wyszecki (1967). The latter experiment was to confirm or modify the findings of Wyszecki and Sanders and involved a total of 76 observers viewing chromatic tiles against a series of achromatic tiles. In analyzing their experimental data in the CIE chromaticity diagram, Wyszecki and Sanders found the resulting iso-brightness to luminance (B/L) contours to vary in irregular fashion (Fig. 5-20). The effect is weakest for yellow and strongest for purple colors. The tile experiment of the OSA committee produced similar contours. Wyszecki and Sanders fitted an analytical formula in the CIE chromaticity diagram that predicts the B/L ratio from the chromaticity coordinates of the color. By that formula the correlation coefficient between experimental and calculated results for their data was found to be 0.92. In 1991 Fairchild and Pirotta found that the magnitude of the implicit additivity failure depends on the luminance level of the surround.

More recently experimental results and interpretations have been reported by Nayatani and co-workers. Their investigations revealed that the magnitude

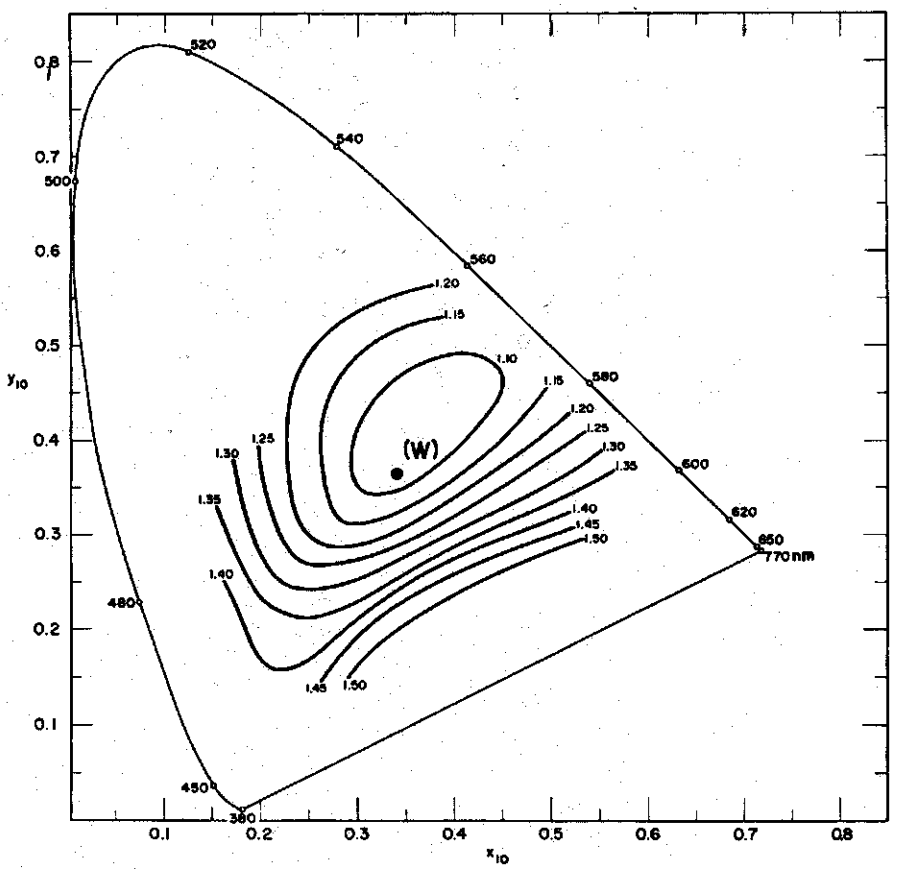

Fig. 5-20 *Iso-brightness to luminance (B/L) contours in the CIE chromaticity diagram as determined by Wyszecki and Sanders (1964), the result of the Helmholtz-Kohlrausch effect. W identifies the white point. The effect is weakest for yellow and strongest for purple colors.*

of HKE depends on the experimental situation (Nayatani et al., 1994a,b). In 1997 Nayatani offered a formula that predicted the Wyszecki-Sanders data with a correlation coefficient of 0.89. It involves the determination of a spectral function that resembles the chromatic threshold function and Evans's function of zero grayness. In 1998 Nayatani proposed a formula where the spectral chromatic strength function was calculated in three sections from Hurvich and Jameson type opponent color functions: below 500 nm it is calculated as the square root of the sum of the squares of the *B* and *R* lobes of the opponent function, between 500 and 580 nm from the *G* lobe, and above 580 nm from the *R* lobe.

In 2000 Kuehni reported analysis results for the Wyszecki-Sanders and the Wyszecki 1967 data (Kuehni, 2000d), using a linear opponent color system. For the former a correlation coefficient of 0.95 was obtained for predicting the B/L

ratios using a formula where the perceived brightness was calculated by adding, by quadrant, a portion of the corresponding linear *a* and/or *b* opponent color value as follows:

$$Y_A = Y_C + 0.23\,|a| \qquad \text{if } b \text{ is positive,}$$

$$Y_A = Y_C + 0.20\,(a^2 + b^2)^{0.5} \qquad \text{if } a \text{ is negative and } b \text{ is negative,}$$

$$Y_A = Y_C + 0.30\,|a| \qquad \text{if } a \text{ is positive and } b \text{ is negative.} \qquad (5\text{-}11)$$

Y_C is the luminosity value of the sample. For all yellowish colors a portion of the positive, respectively negative, lobes of the *a* function is added; for bluish-greenish colors a portion of the geometrically summed two opponent functions is added; for reddish-bluish colors a slightly larger portion of positive lobes of the *a* function is added. To optimally predict the tile data of 1967, larger contributions of opponent color signals in a somewhat different vector composition was required:

$$Y_A = Y_C + 0.43\,(a^2 + b^2)^{0.5} \qquad \text{if } b \text{ is negative,}$$

$$\mathrm{Y}_A = Y_C + 0.28\,(a^2 + b^2)^{0.5} \qquad \text{if } b \text{ is positive.} \qquad (5\text{-}12)$$

This formula predicts the committee data with a correlation coefficient of 0.88. The corresponding effective luminosity function is compared with the CIE luminosity function in Fig. 5-21. When inverted for comparison to the chro-

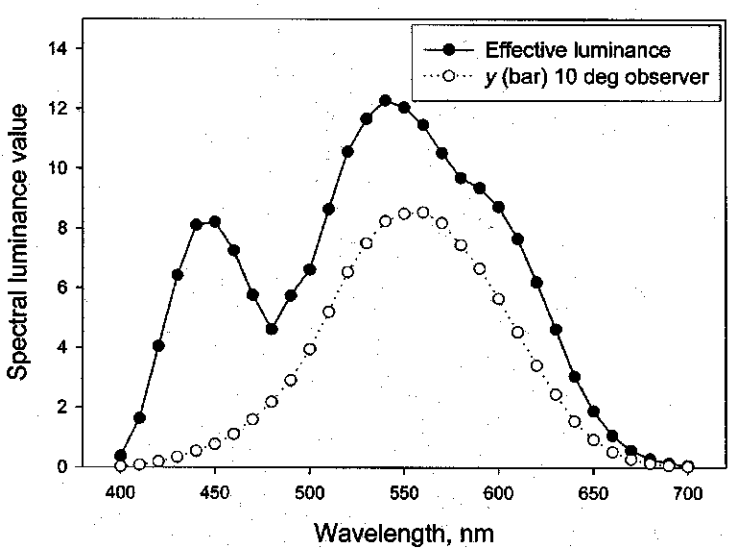

Fig. 5-21 *Spectral effective luminous reflectance for the OSA-UCS Helmholtz-Kohlrausch effect data, as modeled by the author (2000d), compared to the 10° observer ȳ function.*

matic strength function, its long wave trough is found to be shifted toward lower wavelengths. The reason for the difference of the apparent contribution of opponent color signals to perceived brightness in the two experiments is not clear. However, these results provide implicit support for the idea that the opponent color system makes a contribution to perceived brightness/lightness of chromatic colors.

Psychophysical Scaling of Chroma

In their representation in the psychological diagram, Munsell colors of equal chroma are located on concentric circles. In the CIE chromaticity diagram, they lie on ovoid-shaped contours with the major axis in the direction of approximately dominant wavelength 575 nm, meaning in the yellow-blue direction (see Fig. 7-13). There have been many efforts to find models for chroma spacing that are in better agreement with the psychological spacing. A significant improvement is obtained when plotting them in an opponent color diagram. In an Adams type opponent color space, such as the CIELAB space, the contour (see Fig. 5-24a) is improved but continues to be unbalanced in the yellow-blue direction. In 1946 Saunderson and Milner offered an analytical solution modifying the Adams chromatic diagram by changing the scaling of the yellow-blue axis. In an analysis in 1949 by Burnham this adjustment was found to represent the best fit to the Munsell system. However, he offered no explanation of its need.

What Munsell called chroma has been identified in more recent literature as contrast (at constant luminous reflectance). Contrast has been scaled in terms of contrast threshold steps (e.g, see Eskew et al., 1999). But systematic scaling comparable to the psychological scaling discussed below is lacking.

Constant Chroma Circles

As indicated in Chapter 4, since the establishment of the Munsell Renotations only a few systematic determinations of constant chroma around a hue circle have been made. The Re-renotations, a modification of the Renotations for certain findings obtained in connection with the development of the OSA Uniform Color Scales, have chroma scales significantly different from those of the Reenotations, and the constant chroma contour has a different shape and angle of the major axis (see Chapter 7). In 1957–8 Nickerson, Judd, and Nimeroff, also in preparation for OSA-UCS, developed a circle of constant chroma samples at value 6 (Judd, 1965), as mentioned in the previous chapter. It again differs noticeably in shape and angle of the major axis of the ovoid from those of the Munsell system. Chroma scales are also implicit in the OSA-UCS system, assuming that chroma can be calculated as the euclidean sum of g and j. The formulas for g and j represent the committee's euclidean fit to average judgments by 49 to 76 observers of color differences between 128 nearest neighbor pairs of a triangular grid of colors. The resulting ovoid shape

is again significantly different. There is no objective argument for any one of the four scales being most accurate. None of the four psychological chroma scales is well represented by the CIELAB formula. The four "constant" chroma contours are illustrated in Fig. 5-22 a and b in the balanced cone sensitivity opponent diagram [5.67 ($L - 1.9884M$) vs. $L + M - S$], indicating the significant differences. The four scales and the scale implicit in CIELAB have been modeled with cone sensitivity based opponent color functions. The results are listed in Table 5-1, which also contains the coefficient of variation in percent (COV) of the calculated chroma values. The equations indicate that the contours differ primarily in the contribution of the *b* function to the *a* function, namely the implied degree of reappearance of red at the short wave end of the spectrum, and weightings that differ by quadrant. UCS and CIELAB resemble each other in their weightings of the *b* scale, induced by the common cube root power modulation. The fact that the constants for CIELAB are significantly different and the coefficient of variation is much higher indicates that it is not a good representation for any of the four visually based contours.

As we saw in Chapter 4, indirect chroma scales derived from judgments of the content of unique hues in Munsell chips (Indow's work) have not resulted in meaningful scales. As a result we have comparatively poor agreement between various direct and indirect chroma scaling results and no good fit by the formula currently recommended by the CIE. This applies also to the CIE94 formula whose only difference in regard to chroma is that it has an adjustment for chromatic crispening and the basic ovoid shape of the unit difference contour. As will be discussed later, chromatic crispening is likely limited to comparatively small differences.

TABLE 5-1 Cone sensitivity based opponent color models fitted to "uniform" chroma circles

Data	*a*	*b*
Munsell Renotations	$a = 0.926\ [5.67(L - 1.9884M) - 0.50b]$ COV = 5.6%	$b+ = L + M - S$ $b- = 0.90(L + M - S)$
Munsell Re-renotations	$a+ = 0.646\ [5.67(L - 1.9884M) - 0.25b]$ $a- = 0.807\ [5.67(L - 1.9884M) - 0.25b]$ COV = 1.6%	$b+ = 1.09(L + M - S)$ $b- = L + M - S$
Nickerson et al.	$a+ = 0.953\ [5.67(L - 1.9884M) - 0.20b]$ $a- = 1.030\ [5.67(L - 1.9884M) - 0.20b]$ COV = 4.4%	$b = L + M - S$
OSA-UCS	$a+ = 0.908\ [5.67(L - 1.9884M) - 0.40b]$ $a- = 0.976\ [5.67(L - 1.9884M) - 0.40b]$ COV = 4.7%	$b+ = 1.33(L + M - S)$ $b- = 0.75(L + M - S)$
CIELAB	$a+ = 0.676\ [5.67(L - 1.9884M) - 0.25b]$ $a- = 5.67(L - 1.9884M) - 0.25b$ COV = 8.8%	$b+ = 1.33(L + M - S)$ $b- = 0.65(L + M - S)$

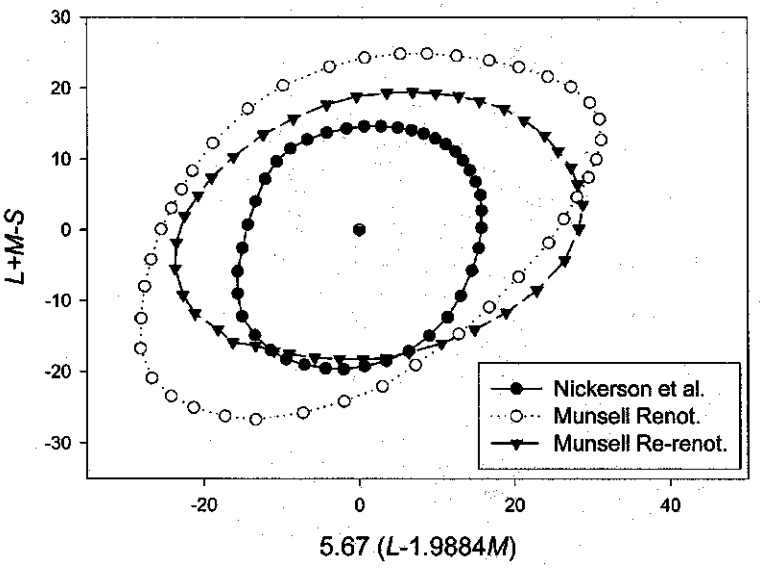

Fig. 5-22a *Constant chroma contours determined in three different experiments in the balanced cone opponent diagram: Nickerson et al. (Judd, 1965), Munsell Renotations, and Munsell Re-renotations.*

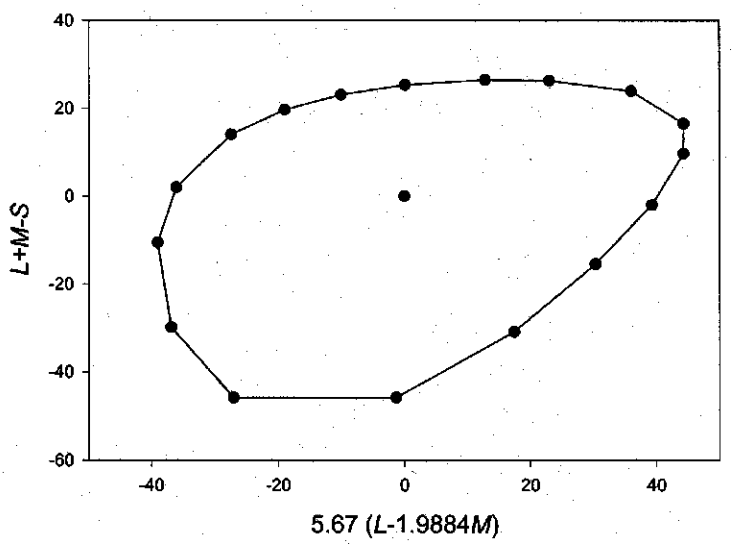

Fig. 5-22b *Constant chroma contour implicit in the OSA-UCS system, in the balanced cone opponent diagram.*

TABLE 5-2 Optimal powers to linearize relations between psychophysical and psychological chroma scale values at or near the axes

Data	*a*+	*a*–	*b*+	*b*–
Munsell Renotations	0.15	0.19	0.42	0.06
COV, %	6.5	4.7	13.1	4.3
Munsell Re-renotations	0.45	0.75	0.70	0.38
COV, %	2.2	6.4	11.1	6.1
OSA-UCS	0.70	0.84	0.58	0.33
COV, %	0.3	0.4	1.8	0.2

Relation of Constant Chroma Circles at Different Chroma Levels

Psychophysical constant chroma contours may be established at different levels of chroma. If they are established so that they are perceptually equally spaced, the question arises as to how the psychophysical values are related for different chroma levels at constant hue and constant luminous reflectance. As was mentioned earlier, an important proposal was made in 1942 by E. Q. Adams. He suggested that the CIE tristimulus values should be scaled in accordance with the Munsell value function. Later the Munsell function was replaced by a power function, and the current CIE recommendation in the CIELAB color space and difference formula employs a cube root function. The same power function was also used in the development of the OSA-UCS formula. However, analysis of the Munsell Renotations, Re-renotations, and OSA-UCS indicates that the effective optimal power is usually different from cube root and differs by semi axis (Kuehni 2000b). Table 5-2 lists the optimal powers to be applied to tristimulus values for the three data sets.

The results, while indicating considerable differences in the optimal powers for the three sets of data, show a degree of relatedness of the Re-renotations to OSA-UCS, as indeed attempted by Judd. The low COV values of OSA-UCS are due to the systematic tiling of the *j, g* diagram once a formula had been fitted to the basic experimental data (see Chapter 7 for details). Given the uncertainty in the related uniform chroma circles and the variability in powers, it is also here not apparent which system, if any, is accurate.

Interrelation between Chroma, Hue, and Lightness Differences

The uncertain relationship between combined hue and chroma differences, on the one hand, and the chromatic and lightness differences, on the other, was discussed in the previous chapter. Color space and difference formulas generally assume euclidean summation. However, the Nickerson and Indow formulas (in Chapter 4) argue for linear addition. The fact that the CIELAB formula is not a good model for any of the experimental chroma and hue scales brings up the issue of conflated component differences calculated from that formula. Figure 5-23 illustrates two colors A and B located on a quarter circle

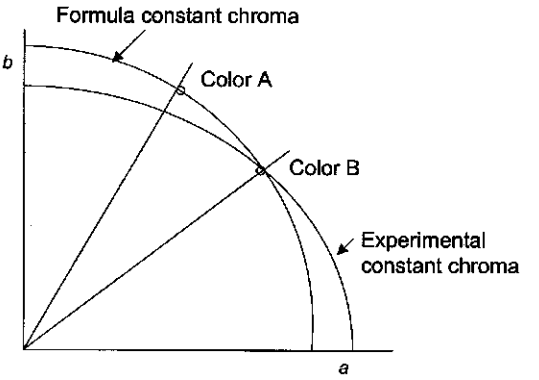

Fig. 5-23 *Schematic depiction of quadrant 1 of an opponent color diagram with constant chroma contours calculated from a formula and determined by an experiment.*

in the a^*, b^* diagram and therefore representing a hue difference only. However, if the true constant chroma contour is represented as indicated in the figure, as might be possible based on the equations in Table 5-1, the difference between A and B is in reality a mixed difference consisting in part of a hue and in part of a chroma difference. The degree of this conflation also depends on the size of the unit hue difference along the constant chroma contour. We cannot expect to have a reliable hue scale unless we have a reliable chroma scale. This fact was recognized by Judd and his co-workers in their work in the late 1950s. A reliable chroma scale, at the same time, depends on a reliable determination of constant perceptual lightness planes.

Additional conflation can result from the different effective powers shown in Table 5-2. It is evident that we cannot rely on hue and chroma differences calculated from CIELAB to have an accurate relationship to the perceptual data of the three data sets. Conflation between lightness and chromatic differences in CIELAB related formulas is also a fact because chromatic and achromatic colors of equal perceived lightness are not on the same value plane due to the Helmholtz-Kohlrausch effect. A further difficulty is that, as will be seen Chapter 7, a euclidean global color space uniform in terms of hue differences and in terms of chroma differences cannot be uniform in both. At the same time it remains to be seen how uniformly and repeatably observers judge component differences in complex small suprathreshold differences. Such experiments are overdue.

The shape of unit difference contours in a uniform color space is by definition spherical. But that space is not automatically euclidean, as will be shown in Chapters 7 and 8. In a euclidean chromatic diagram the unit contour has trapezoid, or perhaps egg-shaped, rather than circular form. In the recent color vision literature various kinds of contours have provided best fits to threshold data in cone contrast space: superellipses, rhombi, and rectangles. For the same

experimental conditions often different contours provided the best fit to data of different observers (e.g., see Sankeralli and Mullen, 1996).

Psychophysical Hue Scaling

The intent of the psychological Munsell chromatic diagram is to be uniform in terms of hue differences. On the other hand, as seen before, in the psychological Munsell diagram unique hues do not fall on diagram axes, and these axes are therefore not psychologically meaningful. It is useful to determine the location of unique hues in the CIELAB diagram to learn if there is a better agreement with the system axes than in the Munsell diagram. This is illustrated in Fig. 5-24a and b, and we find that for the 2° observer two average unique hues fall close to the system axes.

There are significant discrepancies for the green and red unique hues. The CIELAB formula and some formulas recommended later for small color differences imply equal hue angle differences for equal hue differences. CIE94 (see Chapter 6) weights only the magnitude of the calculated hue difference. Various functions adjusting hue differences as a function of hue angle have been introduced into other formulas, including the new CIEDE2000 (see next

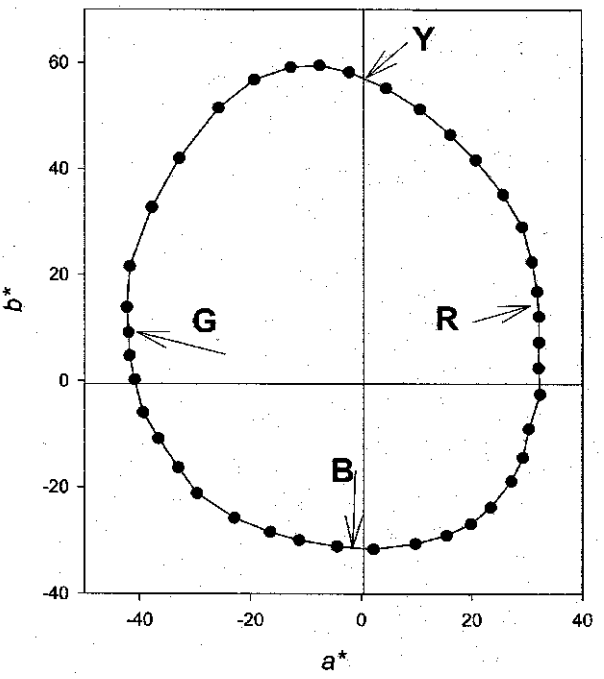

Fig. 5-24a *Munsell Renotation colors of value 6, chroma 8 in the* a*, b* *diagram of the CIE 2° standard observer and the equal energy illuminant. The letters indicate the locations of the average unique hues determined in the experiment by the author (2001a).*

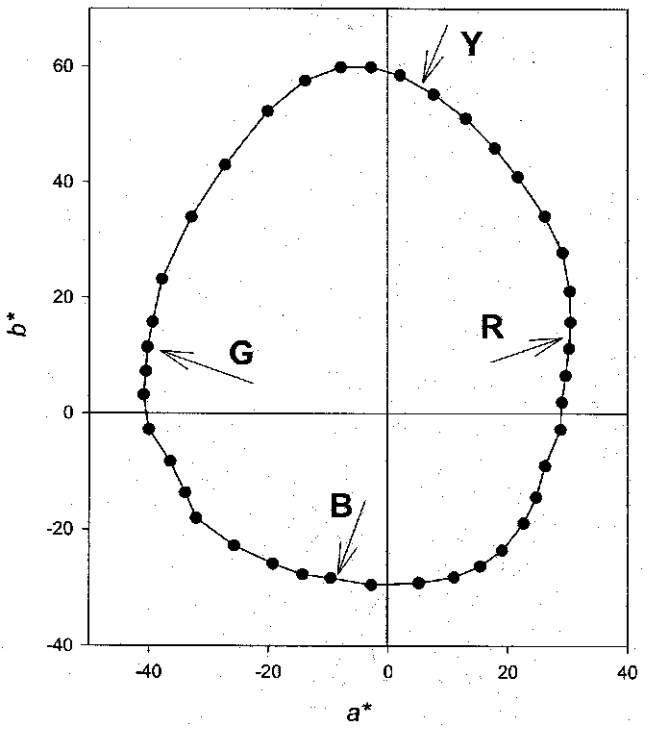

Fig. 5-24b *Colors of Fig. 5-24a illustrated in the* a*, b* *diagram of the CIE 10° standard observer, with locations of average unique hues.*

chapter). These functions not only embody the fundamental variability (as related to uniform hue difference) in hue angle difference but also correction for the error added by the CIELAB formula.

We can surmise from Table 4-1 that the various described experiments have somewhat different hue scales. We find this to be the case when comparing hue angle differences in the optimized cone sensitivity based opponent diagrams of Table 5-1 around the hue circuit for the Munsell Renotations, the Re-renotations and the Newhall data. These are the only data sets where we have continuous consistent hue scaling around a psychologically scaled constant chroma circle (see Fig. 5-25 for the Newhall data). The results indicate little or no discernible agreement except for the region of hue angles 0 to 100 for the Newhall and Re-renotation data. While the changes in hue angle difference appear to be random in the Renotations, there are systematic minima near hue angles 75° and 275° in the Re-renotations. The Newhall data have a systematic development in roughly the first half of the hue circuit. Overall, the lowest level of disagreement is between the Renotations and the Newhall data. All three sets of hue angle differences have slightly positive slopes relative to the hue angle—Renotations 0.0099, Newhall 0.0076, Re-renotations 0.0048—indicating a tendency for the increment to become larger as the hue angle increases.

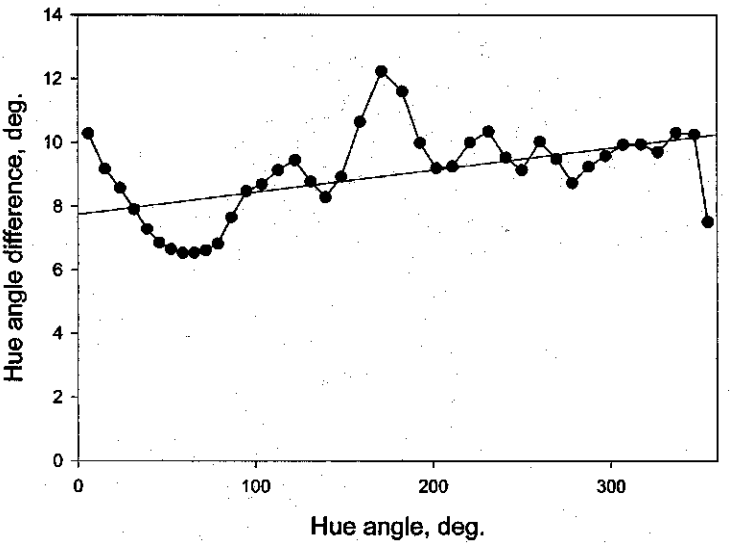

Fig. 5-25 *Hue angle differences in degrees as a function of hue angle between uniform hue steps on a constant chroma contour in the optimized cone sensitivity based opponent diagram, Newhall data. The straight line is a linear regression line.*

Partial scaling of hue circles has been done by Qiao et al. (1998) and several color vision scientists (e.g., see Krauskopf and Gegenfurtner, 1992). It is also implicit in small color difference data (but assuming that a total color difference can be split into components based on the euclidean model and that the CIELAB formula describes visually uniform chroma circles). Differences in stimulus increments needed at different hue angles may be the result of different specific mechanisms being responsible for hue identification and discrimination.

5.9 COLOR MATCHING AND APPEARANCE SCALING

Color-matching functions are expressive of the response of the visual system to the color-matching task. This task does not necessarily have a direct connection to color appearance scaling, even though both presumably are derived from the cone responses. The fact that CIE opponent color diagrams roughly match a Hering type psychological diagram has been suggestive of such a connection.

As mentioned, opponent axes for the 2° observer differ from those for the 10° observer. The differences between the 2° and the 10° color-matching functions are thought to be primarily due to the effect of the macula in the eye (e.g., see Oleari, 1999). In addition the distribution of cone types differs as a function of field size, and the optical density of the cones may change (Pokorny et al., 1976). Macula is a layer of matter containing yellow-colored carotenoids and other compounds. It may have several purposes, one being protecting cell layers

of the more central area of the retina by absorbing short wavelength light. The macula has the form of an irregular spot centered on the fovea, the most sensitive region of the retina. However, the central region of the fovea is free of macular material. There is the interesting situation of the central fovea believed to have no *S* cones and being macula free.[10] Informal tests, with rare exceptions, have not resulted in any discernible hue differences when simultaneously viewing the same colored paper with a gray mask with 2° and 10° openings. Small hue changes have been noted in case of some complex grays and some purples. Chroma is often slightly enhanced for small fields compared to larger ones, perhaps as a result of increased contrast at the lower field size.[11] The 2° observer geometry results in the image being focused to a good extent on the macula-free central fovea. The 10° observer geometry, on the other hand, results in a considerable filtering effect by the macula expressed in the different functions. Nevertheless, appearance of colored materials against a neutral surround seems largely unaffected by field size from 2° to 10°.

From the differences between the two sets of observer functions one can predict that metameric matches viewed in a 2° field generally will not match in a 10° field. In regard to appearance, the "blue" system axis falls on dominant wavelength approximately 477 nm for the 2° observer and approximately 470 nm for the 10° observer. It means that if unique hue for the 2° observer falls on the negative *b* axis, that axis for the 10° observer represents a slightly reddish blue. For the "yellow" axis the values are 578 nm, respectively 573 nm, implying in case of the 10° observer a slightly greenish yellow (as demonstrated in the OSA-UCS system). The 10° blue-yellow axis can be rotated by a linear transformation so that it is much closer to the perceptual unique blue and yellow. The $\bar{x}$ modification required is approximately as follows:

$$\bar{x}_{10,\mathrm{mod}} = 0.9\bar{x}_{10} + 0.1\bar{z}_{10}. \qquad (5\text{-}13).$$

However, such a change also results in significant changes in predicted chroma of a hue circle. The difference between the two CIE standard observers therefore is not just a matter of axis rotation but represents a more fundamental change. It is evident that very specific forms of color-matching functions would be required to reduce reflectances to geometric patterns that closely resemble the geometrical pattern of the psychological space. However, it is questionable if such a transformation exists. There is evidence, to be discussed later, that such a simple approach of combining facts of color matching with facts of color appearance is insufficient. Thus appearance is not in good agreement with the changes implied by the color-matching functions.

5.10 PLACEMENT OF THE RED AND GREEN UNIQUE HUES IN THE OPPONENT COLOR DIAGRAM

The dominant wavelengths of objects seen as having unique red hue are not spectral; their complementary wavelengths fall for most observers on the

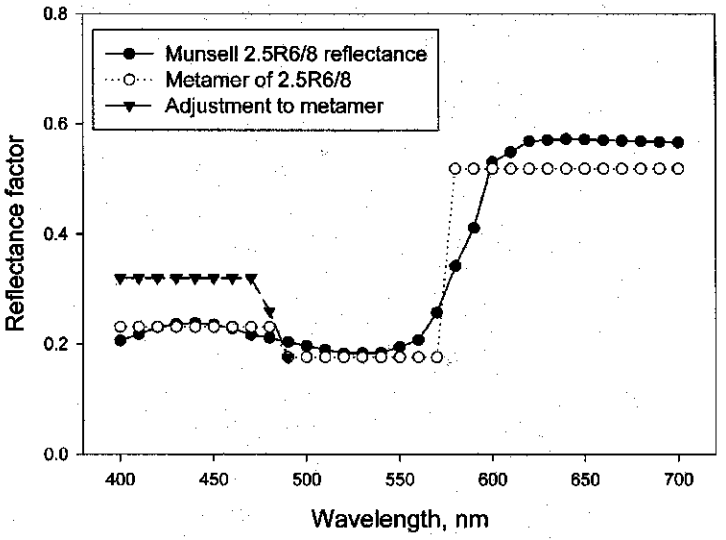

Fig. 5-26 *Spectral reflectance functions of a Munsell color chip (2.5R6/8) nearly representing average unique red hue and a modified red. Open circles represent a simplified metamer of 2.5R6/8. Inverted triangles show the short waveband adjustment required to place the color on the positive* a* *axis of the CIELAB chromatic diagram.*

purple line in the CIE chromaticity diagram. This is indicated by the fact that objects with red appearance have dominant wavelengths not only in the long wave but also in the shortwave region. Conceptually unique red is located where blueness and yellowness are in balance. The reflectance function of an object resulting in unique red hue appearance in standard daylight against an achromatic surround has a small short wave and a larger long wave loop. The real and an idealized reflectance curve representing Munsell color 2.5R6/8, closest in hue at that level of chroma and value to average unique red, is illustrated in Fig. 5-26. While for the 2° observer the tristimulus values are identical, there is a calculated color difference of 0.9 CIELAB units between them for the 10° observer. When calculating the position of this hue in the a^*, b^* opponent color diagram, we find that its hue angle is not near 0° but at 18° (15° for the 10° observer). The Munsell hue located at 0° for both observers is 5RP, for most observers a distinctly bluish red. Instead of near zero, the b^* value for 2.5R6/8is 12.1, indicating that there is too much implied yellowness or not enough blueness present to balance the two. One can determine how much higher the shortwave reflectance would have to be (illustrated in Fig. 5-26) to result in a b^* value of zero. It is relatively a 40% increase. It is possible to shift the weight of the b+ opponent function on the spectral scale while maintaining its total weight in such a way that b^* has a considerably smaller value. But as a result dramatic imbalances of the chromas of a Munsell hue circle are obtained. A similar situation applies to unique green. In fact there

does not appear to be a form of color-matching functions that causes objects of all four unique hues to fall on the system axes while maintaining near constant chroma for the hue circle. This situation may point to an ultimate incommensurability between color matching and color appearance. The implication is that opponent signals of the type recorded in the LGN and visual area V4 of the brain undergo additional nonlinear changes in the "black box" between area V4 and the appearance of colors in our consciousness.

5.11 CURVATURE OF LINES OF CONSTANT HUE BLUE COLORS

Another indication of the discrepancy between color matching and color appearance is the curvature in *a, b* diagrams of lines of constant hue of bluish colors. The curvature is evident when plotting the near-unique hue 5PB at value 4 and several chroma levels in the *a*, b** diagram. It is even more distinctly noticeable when plotting the bluish colors of OSA-UCS in the 10° observer based *a*, b** diagram. Further indication is the rotation of unit color difference ellipses of bluish colors in counterclockwise direction in the *a**, *b** diagram (Fig. 5-27). This effect can be fixed mathematically by using a hue angle based function (see Luo, Cui and Rigg, 2001). The optimal function has to be different for the two CIE observers.

A clockwise rotation takes place when linearly transforming the $\bar{x}$ color-matching function by subtracting a small amount of $\bar{z}$ and replacing it with the corresponding amount of the original $\bar{x}$:

$$\bar{x}_{\text{mod}} = (\bar{x} + m\bar{x}) - m\bar{z}, \tag{5-14}$$

where *m* is a factor with a value of 0.15 or less and depends on the visual angle occupied by the sample; that is, the optimal value for *m* differs for the two CIE standard observers (Kuehni, 1999). The effect of this equation is to adjust the relative size of the two loops of the $\bar{x}$ function by slightly reducing the short-wave and enlarging the long wave loop. We have seen above that adjustments of this type (expressed in terms of cone sensitivity) are also important for optimal fitting of constant chroma data. The effect is also evident in the short-wave red loop of the OSA-UCS *g* function (see Fig. 7-17b). P. W. Trezona and R. P. Parkins (1998) pointed out that the size of the two loops in the CIE color-matching functions is somewhat arbitrary. Application of the revised function rotates not only bluish but also (if less so) yellowish colors in the *a*, b** diagram.

5.12 MUNSELL COLORS IN THE *L, M, S* AND *X, Y, Z* SPACES AND THE A, B DIAGRAM

For a better understanding of how the three psychophysical color spaces are related, it is instructive to compare a selection of Munsell colors (Kuehni,

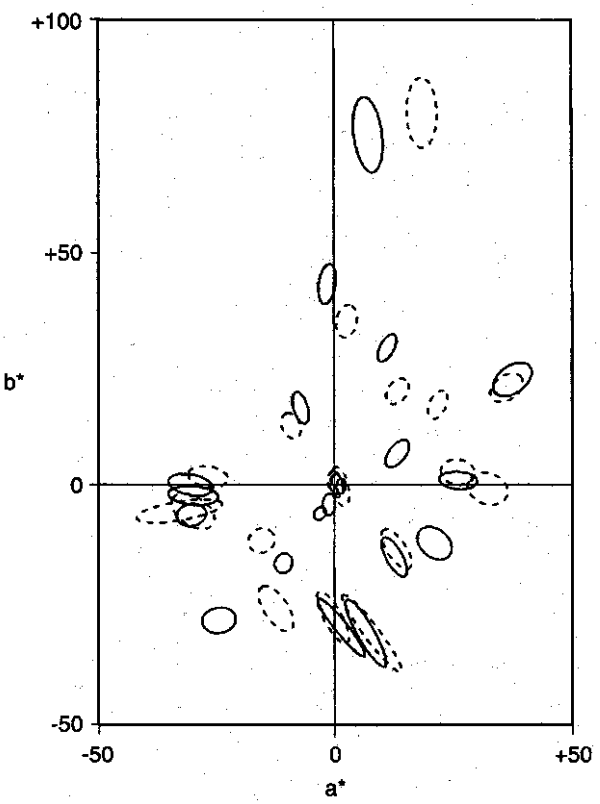

Fig. 5-27 *Ellipses in the* a*, b* *diagram fitted to selected Luo and Rigg (*solid*) and to the RIT-DuPont (*dashed*) small color difference data. Note the rotation of ellipses near the negative* b* *axis. From Melgosa et al. (1997).*

2001c). For this purpose a "Celtic cross" figure is selected consisting of a hue circle at chroma 8 and colors at increasing chroma from the neutral point of hues falling nearest the axes of the *a, b* diagram at values 3, 6, and 8. It is understood that the hue and chroma scaling of these colors cannot be taken as perfect, but there is no doubt that the corresponding samples are a reasonable approximation of a uniform hue circle and uniform chroma scales. In the psychological chromatic diagram the three Celtic crosses coincide forming a single cross. Figure 5-28 illustrates these colors in the three-dimensional psychological diagram. After converting the CIE tristimulus values (adjusted to reflect an equal energy illuminant) to *L, M, S* values using the Smith-Pokorny transformation, the colors are plotted in the *L, M, S* space in Fig. 5-29. The three crosses form sections through an elliptical funnel. Circles in the perceptual diagram are strongly elongated along the *S* axis in the *L, S* and *M, S* planes. The chroma steps along the *S* axis become increasingly larger as *S* increases. Because of the definition of *Y* in terms of *L* and *M* the slices of the elliptical

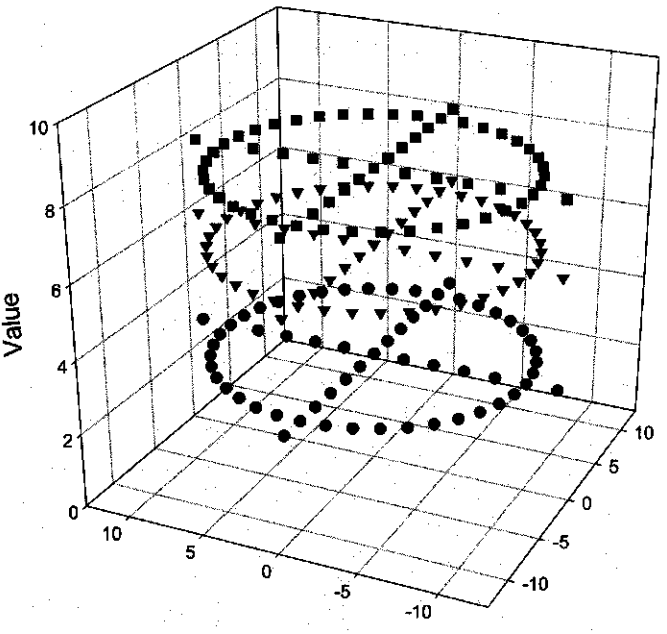

Fig. 5-28 *Representation of Munsell colors at values 3, 6, and 8 and chroma 8, and chroma scales on the axes (Celtic crosses) in the conceptual Munsell color space where hue, value, and chroma scales are not matched.*

funnel are not perpendicular to the equal energy axis but slanted. It is evident that a series of constant value/constant chroma Munsell hue circles has a complex form in the *L, M, S* space without a simple relationship to any of the three dimensions.

The situation improves when plotting the same colors in the *X, Y, Z* space (Fig. 5-30). Since in this space one of the dimensions is luminous reflectance, colors of a given Munsell value fall on planes perpendicular to the *Y* axis. The chroma axes of the cross are well aligned with the *X* and *Z* axes of the space. Scaling of *X* against *Z* is required to make the contours circular. This is evident if we plot the three crosses in a normalized linear opponent color diagram where $a = 2.272\ (X - Y)$ and $b = Y - Z$ (Fig. 5-31). The funnel of hue circles is now parallel to the *Y* axis. In order to convert the funnel to a cylinder in agreement with the psychological cylinder and at the same time create uniform chroma scales, we must determine the applicable psychophysical function(s). As mentioned above, in 1942 Adams proposed that the Munsell value function was applicable to all three tristimulus values, and this proposal has been enshrined in the CIELAB formula. However, the three Munsell Celtic crosses require power modulation that is different for the four semi axes (Table 5-2), and additional modulation to convert the cone in tristimulus space with good accuracy into a cylinder (more on this subject in Chapter 7)

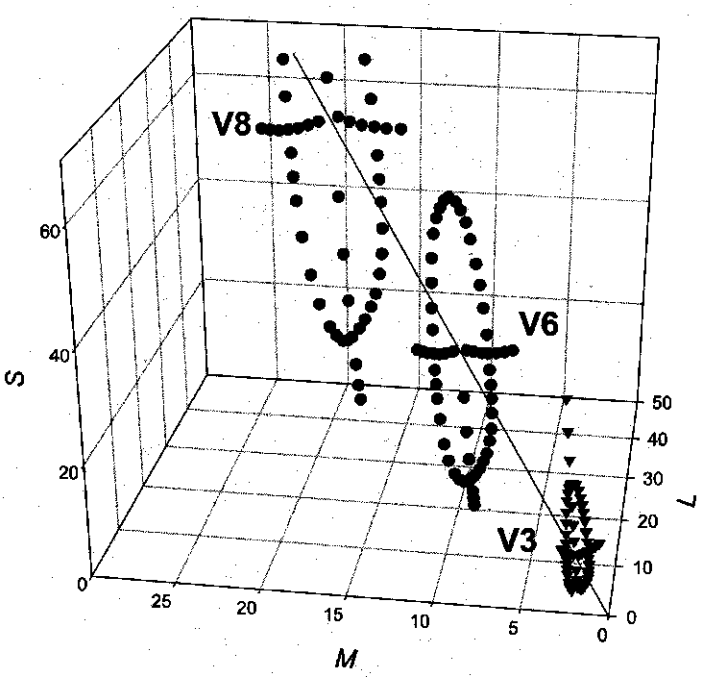

Fig. 5-29 *Actual Munsell colors approximating the Celtic crosses of Fig. 5-28 in the* L, M, S *cone sensitivity space; 2° observer, equal energy illuminant. The straight line represents the equal energy locus.*

5.13 SUPRATHRESHOLD SMALL COLOR DIFFERENCES

Differences in all sets of such data, as will be seen in the next chapter, have been judged as total differences. In nearly all cases the total differences, as calculated by the CIELAB formula, involve differences in all three attributes. The historical process of fitting psychophysical formulas to such data is described in Chapter 6. CIELAB did not result in good correlation against visual judgments in various data sets, and efforts since the mid-1970s usually involved modification of CIELAB to improve correlation. The most widely used formulas today (in the United States) are CMC and CIELAB. The CIE has issued recommendations of new formulas in 1994 and 2001. As in the case of hue and chroma scaling data sets, different sets of small color difference judgments vary considerably. A formula fitted to one set of data often does not fit another set well. Some of the reasons may have to do with different experimental conditions, often inadequately described. Other reasons may involve the composition of the observer pools used in the experiments and, possibly, cognitive components in the judgments of the observers. Clarification of the reasons behind the psychological variability is important for accurate model building.

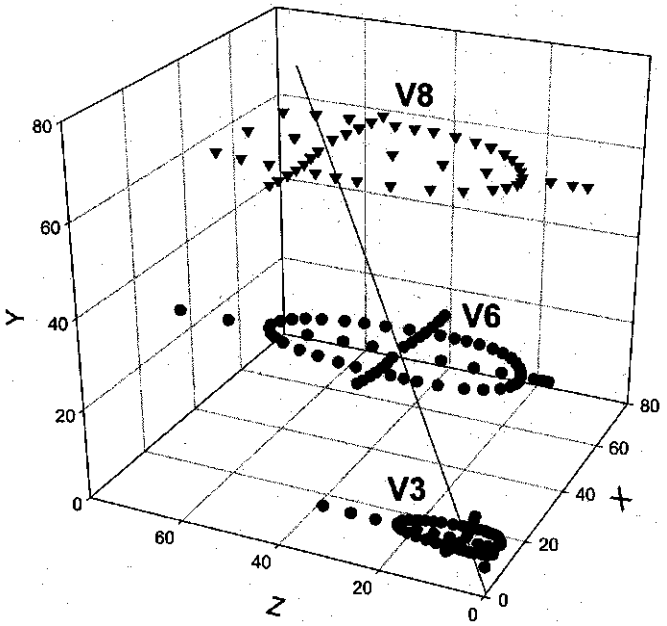

Fig. 5-30 *Colors of Fig. 5-29 in the 2° observer* X, Y, Z *tristimulus space, equal energy illuminant. The straight line represents the equal energy locus.*

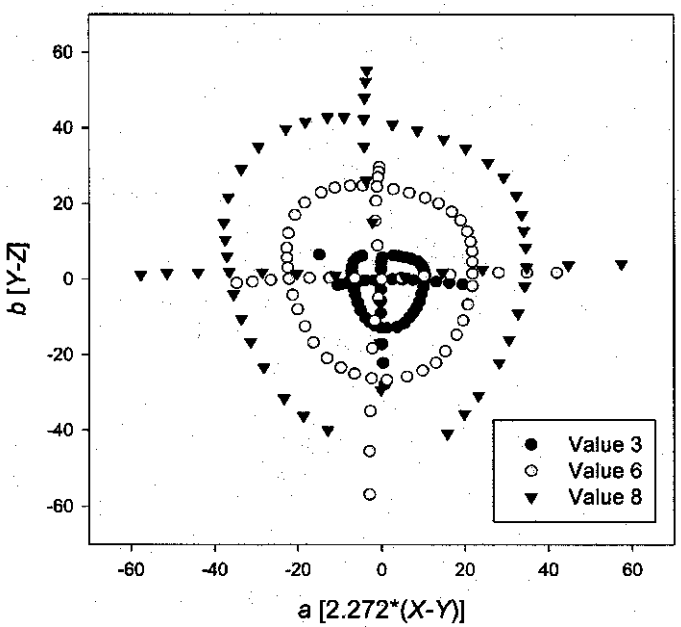

Fig. 5-31 *Colors of Fig. 5-29 projected onto a balanced linear opponent color diagram based on tristimulus values (2° observer, equal energy illuminant).*

5.14 DIFFERENCE THRESHOLD MEASUREMENTS

As will be discussed in the next chapter early measurements of color discrimination thresholds expressed in psychophysical systems were made by Wright (1941) and by Stiles (1946). Newer measurements important for the purposes of this book, using sophisticated computer control of visual displays with tight controls of adaptation, are those by Krauskopf and Gegenfurtner (1992), Strohmeyer et al. (1998), Sankeralli and Mullen (1999), and Smith, Pokorny, and Sun (2000). In recent years threshold measurements using object color samples have been made by Richter (1985) and by Witt (1987).

Typical measurements in vision science involve the so-called pedestal technique. A color stimulus is displayed briefly (one second or less) against a surround that controls adaptation. A small stimulus, typically at constant luminance, is added to or subtracted from the base pedestal stimulus and randomly displayed against it. The observer indicates if she detected a difference between the two displayed fields. The results are now generally reported in a Derrington, Krauskopf, and Lennie type chromatic diagram (see Section 6.17) with its axes scaled in terms of the threshold at surround chromaticity. Among the many results are the following:

1. Thresholds are always smallest in terms of incremental stimulus when the stimuli being compared have luminances or chromaticities straddling those of the surround. This is confirmation of the results of Schönfelder (1933).
2. When measuring chromatic thresholds against a neutral surround, the stimulus increment required for a criterion response increases with increasing contrast between surround and the reference pedestal, except along the system axes.

In this kind of experiment only increments along the $(L + M) - S$ axis increase, while those along the $L - M$ axis do not. Close to surround the increment actually slightly declines in size compared to that exactly at surround (Sankeralli and Mullen). Threshold contours fitted to color centers at some distance from the surround chromaticity are ellipselike, pointing in the general direction of the diagram origin (Krauskopf and Gegenfurtner, 1992; see Fig. 8-14). Such findings have resulted in the idea of multiple hue detection mechanisms throughout the chromatic plane, a crucial issue in regard to modeling of hue discrimination. Mechanisms of such a type are believed to be of cortical nature, but no specific neurophysiological mechanism has been proposed yet. The idea is controversial as seen from arguments by Krauskopf (1999) and by Eskew et al. (1999).

The outcome may depend on the outcome of another controversy. Derrington et al. (1982) found that most parvo cells in the lateral geniculate nucleus are tuned to their cardinal chromatic axes (not unique hues), indicating no hue angle dependent selectivity. Later, in visual area V1 of the cortex,

such clear orientation of cell response is absent, even though there is a slight tendency in this direction (Lennie, 1999). The same applies to areas V2 to V4. De Valois et al. (1997) found that LGN cells of all types have input into later cells with color sensitivity. This may explain the differences in threshold measurement results along the cardinal axes compared to away from them. These results point to the possibility of a rather complex hue detection mechanism that may depend on averaging computations among many cells with specific hue tuning.

5.15 HOW MANY COLORS CAN WE DISTINGUISH?

This is an interesting question for which several answers have been given in the past. The first determinations of this kind addressed the limited issue of the number of discriminable steps in the spectrum. Initial experimental attempts to answer this question were made by E. Mandelstamm, reported in 1867. Measurements considered valid for many years were those by W. Dobrowolsky (1872). Kries computed in 1882 from Dobrowolsky's data 208 discriminably different steps. Based on König and Dieterici's experiments of 1884 the number increased to 235. However, there is the issue of conflation of hue, saturation, and brightness differences when looking at spectral colors. In addition the results depend on the experimental setup. When correcting for brightness differences, L. A. Jones in 1917 found 128 hue steps. Many determinations of wavelength discrimination have been made since then, notably by Wright and Pitt in 1934 and Bedford and Wyszecki in 1958, with quite similar results. MacAdam (1947), on basis of his color-matching error data, calculated a number of 250 just noticeable differences in the spectrum, not far different from that of König and Dieterici.

But color, as we know, is not just hue but also brightness/lightness and saturation/chroma. In 1896 Titchener estimated the total number of perceptibly different colors at about 33,000. By 1939 the number, as estimated by Boring, had risen to 300,000. Both Titchener and Boring did not distinguish between light colors and object colors. In the same year Judd (1939) estimated the number of perceptible object colors as 10 million. This figure has become much quoted. In 1943, in connection with their development of a psychological color solid (see Figs. 1-2 and 1-3), Nickerson and Newhall calculated a rounded number of 7,500,000 at the just noticeable difference level (applicable to the model on the right side of Fig. 1-2) and a number of 1,875,000 under less favorable viewing conditions (applicable to the model on the left side). This number applies to a color solid reaching out to the object color limits.

MacAdam in 1947, based on his color matching error data of 1942, limited himself to determining the number of perceptually different colors in a constant luminance plane. Using three standard deviations of the matching error as the JND limit, he calculated the number of colors distinguishable in a constant luminance plane, determined under the conditions of his color-matching error experiment, in a rounded number, as 17,000.

More recently the question has appeared again. In 1981 A. Hård and L. Sivik (in contrast to the 10 million colors Judd mentioned as discriminable) estimated the number of distinct colors that can be identified with a degree of certainty as 10 to 20,000. In 1995 Indow estimated the number of colors that can be discriminated from each other as approximately 7 million without providing an explanation for the estimate. In 1998, using the CIELAB formula as a basis and assuming the limit of object color distinction to be 1 CIELAB unit of total color difference, M. R. Pointer and G. G. Attridge calculated the number of distinguishable object colors to be 2.28 million. They also calculated the theoretical limit of different object colors displayable on a video monitor to be 16.78 million.

Based on a single medium gray background/surround the number must be significantly smaller than 2.3 million because of lightness and chromatic crispening effects. At the same time the JND limit near the surround is likely smaller than 1 CIELAB unit. The CIE94 color difference formula predicts that one CIELAB chroma unit at metric chroma zero is only approximately 0.2 chroma units at metric chroma 100. Unless there is a widely variable surround the number of perceptibly different object colors is more likely about 1 million. Nevertheless, it is a remarkable number of different percepts based on the output from three cone types.

There is, however, an additional issue. The number of 1 million or 2.3 million applies to the situation of placing a net defining JNDs over a much finer net defining visible stimuli. Once a starting position is defined, one can perceive 1 million other colors compared to the reference color. But the reference color can be varied within the just noticeable difference limen around it. Applying the JND net to a new reference point, we experience 1 million slightly different colors. The number of colors we can experience may, after all, be more in the range of what can be produced on a video monitor.

In this chapter it has been shown that the fitting of psychophysical models to uncertain psychological data has added another layer of problems. In the absence of full understanding of the human visual mechanism, such modeling is mainly empirical. The currently neurophysiologically supported (in macaque) opponent space at the LGN level is not in agreement with the human psychological opponent color space. But a comparatively simple adjustment can improve the fit. Interestingly different experimental psychological constant chroma circles can be modeled closely with only changes in the level of this adjustment and modifications in scaling of the semi axes. It is necessary to investigate the causes of differences in constant chroma evaluations: uncontrolled surround effects, variability in the personal color-matching functions of the observer pool, perhaps other reasons. Different sets of hue scaling data around a constant chroma circle also have provided significantly different results. In both cases development of reliable, replicated data seems urgent.