Factors underlying individual differences in the color matches of normal observers

Michael A. Webster*

Department of Psychology, University of California, Berkeley, Berkeley, California 94720

Donald I. A. MacLeod

Department of Psychology, University of California, San Diego, La Jolla, California 92093

Received March 13, 1987; accepted May 31, 1988

We have used a factor analysis of the Stiles–Burch [Opt. Acta 6, 1 (1959)] 10° field color matches to examine the basis of individual differences in the color matches made by observers with normal color vision. The differences in the matches are primarily due to interobserver variations in the macular-pigment density [with a standard deviation (σ) of 0.12 at 460 nm]; the lens-pigment density (σ = 0.18 at 400 nm); the spectral position of the long-wavelength-sensitive (σ = 50.3 cm^{-1}), medium-wavelength sensitive (σ = 31.9 cm^{-1}), and short-wavelength-sensitive (σ = 45.3 cm^{-1}) photopigments; the covarying densities of the three photopigments (σ = 0.045); and the degree of rod intrusion. Variations in the different factors appear to be uncorrelated. Comparable estimates of the sources and range of interobserver differences in color matching were obtained from a similar analysis of the Stiles–Burch 2° color matches [Opt. Acta 2, 168 (1955)].

INTRODUCTION

Observers with normal color vision often make substantially different color matches. One well-known reason for these differences is that individuals vary widely in the amounts of the prereceptoral screening pigments present, such as the lens and the macular pigments. For example, the lens pigment absorbs a large fraction of the short-wavelength light entering the eye but little of the longer wavelengths. Thus an observer with a higher than average concentration of this pigment will require a less than average intensity of a red primary light when matching a blue test light, simply because the blue light has been selectively more attenuated. By expressing the color matches in a form such as the WDW coordinate system, the effect of any inert pigment screening all the receptors is eliminated.[1] Yet substantial individual differences in the matches remain, suggesting that the photopigments and/or receptor mechanisms also vary among observers.

It is thus clear that the interobserver differences in color matching are due to a large number of factors. However, exactly what this set of factors is, and how much each contributes to the total observed variability, is not definitely known. In this study we have examined these questions by parceling out and identifying the sources of variability in a set of normal color matches with the technique of factor analysis. The analysis not only allows us to determine which potential factors influence the matches and the range over which they vary but also allows us to examine such properties of the factors as the spectral absorption characteristics of the pigments involved.

The different steps of our analysis are organized as follows. First, we use standard factor analytic techniques to estimate the underlying patterns of individual variation in the Stiles–Burch 10° field color matches.[2] We then identify the variable physiological characteristics to which these observed factors correspond and determine how much these characteristics vary across observers. We also examine whether the variations in the different factors are independent or correlated. Next, we discuss a theoretically guided factor analysis, which was designed to improve our estimates of the pattern of variability due to each factor, by taking advantage of some established properties of color matching. Following this, we derive separate estimates of the range of interobserver differences in the identified factors, by finding the amount of variation in each predicted factor that provides the best fit to the observed correlations between the different color matches. Finally, we perform a similar analysis of the Stiles–Burch 2° color-matching data.[3] In the accompanying paper,[4] we use individual differences in the color matches, and independent data on the changes in color matches at high light levels, as a way to estimate directly the absorption spectra of the cone photopigments.[5]

METHODS

In the standard color-matching experiment the stimulus consists of a circular split field. On one side a test light of a selected wavelength and fixed radiance is presented, and this is matched by a mixture of three primary lights of suitably chosen, fixed wavelengths. The match is achieved by adjusting the radiances of the three primaries until the two half-fields appear equal in both color and brightness. For spectral test lights this is generally possible only by adding one of the primaries to the test half-field, in which case the sign of the radiance of that primary in the matching mixture is considered to be negative.

The particular data that we chose to examine were the 10° field color matches of Stiles and Burch for 49 observers.[2] These represent the largest comprehensive study of normal

0740-3232/88/101722-14$02.00

color matching. Measurements were made for 32 monochromatic test lights ranging from 392 to 714 nm, in equal-wave-number steps of 250 cm^{-1} between 455 and 625 nm and in steps of 500 cm^{-1} outside this range.[6] The results were reported in terms of red-, green-, and blue-matching primaries at 645, 526, and 444 nm, respectively.[7]

For our purposes, one can think of this experiment as the measurement for each observer of 96 variables, namely, the logs of the radiances of each of the three primaries required to match each of the 32 test wavelengths. Not surprisingly, many of these variables are redundant. For example, the radiances of the 444-nm primary required to match test lights of 400 and 408 nm are highly correlated with each other—an individual who requires a lower than average radiance to match one will require less than average for the other. Presumably these relationships between the observed variables are due to the common influence of a smaller set of underlying variables, or factors. Thus the amounts of the 444-nm primary in the 400- and 408-nm matches are not each measures of two separate determinants of the color matches but rather are assumed to be largely two measures of the same influences.

Given this assumption, the goal of factor analysis is to use measurements of the observed variables in a set of individuals to define the underlying variables and their values for each individual. This is done by examining the correlations among the observed variables. Correlation among a group of variables implies the existence of one or more common underlying factors. The degree to which any single observed variable is influenced by a factor is measured by the correlation between them, termed the factor loading. If the factor loading for some variable on a factor is 1.0, then the two are perfectly correlated and all the variable's variance is due to variations in that factor alone. More typically, though, values of an observed variable will reflect the influence of more than one factor and will thus have loadings of less than 1.0 on two or more factors. In standardized form, each measured value z_{jk} of an observed variable j in an individual k is regarded as the result of contributions from each of the n underlying common factors, and these contributions in turn depend on both the value F_{ik} for factor i in the individual case and the factor loading a_{ji} expressing the influence of factor i on variable j:

$$z_{jk} = a_{j1}F_{1k} + a_{j2}F_{2k} + \ldots a_{jn}F_{nk} + a_{ju}U_{jk}.$$

U_{jk} is a factor corresponding to any variance that is unique to j. Note that the factors do not provide any information about the actual values of the color-matching functions but depend only on the variations in those functions across observers, as reflected in the correlation matrix. The squared loading of any variable on a factor represents the proportion of the total variance of the variable that is due to variations in that factor. Thus the contribution of a factor to the total standardized variance in the data set is given by the sum of its squared loadings. For the single variable j, the sum of the squared loadings across the n common factors,

$$h_j^2 = \sum_{i=1}^{n} a_{ji}^2,$$

is termed the communality and is equal to the proportion of observed variance accounted for by the set of common factors.[8]

The dependence of an observed variable on the common factors can be visualized by representing it as a point in a multidimensional space, where the axes of the space correspond to the different factors and the point's coordinates are its factor loadings. Mathematically, this is simply an alternative representation of the original correlation matrix—within this multidimensional space the correlation between any two observed variables is given by the cosine of the angle that they subtend at the origin (times the product of the two vector lengths from the origin, which represent the communalities for each variable). However, to the extent that the number of factors is small relative to the number of observed variables, this new representation has the advantage of describing the data much more parsimoniously, in terms of a small set of dimensions. In our case these dimensions should represent those properties of the visual system that underlie the individual differences in the color-matching functions defined by the 96 measurements on our 49 subjects, and the method thus allows us to specify what those properties are and over what range they vary. The procedure is particularly well suited to the analysis of color matches, both because of the high reliability with which the matches can be made[2] and because the significant underlying variables are probably few in number and well defined by the observed measurements.

A serious problem with this representation, however, is that it is not unique. Rotating any of the axes through any arbitrary angle produces a mathematically equivalent solution, while resulting in a radically different pattern of factor loadings—one that might lead to a quite different theoretical interpretation. Suppose, for instance, that two uncorrelated variables in a two-dimensional factor space had loadings of (0, 1.0) and (1.0, 0). This would suggest that each of the variables is influenced by only one of the factors. However, we could just as well have described the relationship between the observed variables by rotating the axes 45°, in which case the angle between the variables would remain 90° but the loadings would change to (0.707, 0.707) and (−0.707, 0.707). Yet this representation would suggest that the underlying dimensions have the property that both influence both variables to the same extent (though for one of the factors in opposite ways). Obviously the identity of the factors inferred from these two rotations would be very different.

The interpretation of the extracted factors thus depends critically on the orientation of the axes, and some assumptions must therefore be brought to the analysis to constrain the rotation. Initially we relied on the general Varimax rotation criterion. This makes no specific assumptions about the nature of the underlying factors but assumes only that different variables tend to be influenced by different factors and that the underlying factors are orthogonal. The rotation is then found that best satisfies these criteria within the constraints of the data; specifically, this is done by finding the solution that maximizes the variance in the squared loadings for each factor. (Note that in the example above the squared loadings for the first rotation have the maximum possible variance, while for the second it is at a minimum; the Varimax rotation would therefore converge on the first.) As we shall see, this one general assumption was for

the most part sufficient to guide our analysis to the theoretically correct rotation.

RESULTS

Initial Factor Analysis

The basic factor analysis was performed with the Statistical Package for the Social Sciences (SPSS) subprogram FACTOR[9] by inputting 97 items for each subject (the 96 primary settings plus the subject's age). Initial factors were extracted from the item correlation matrix by using a procedure that iteratively estimates the communalities, which determine how much of the variance is to be partitioned into common factors, and then derives factors based on those estimates.[10] The number of factors determined in this way was originally limited to 20, with each successive factor accounting for a smaller proportion of the variance. However, only a subset of these presumably corresponds to real properties of the visual system, as opposed to random noise in the measurements, and these could be qualitatively identified by a pattern of loadings that varied systematically with wavelength. The first 10 factors clearly met this criterion, and the first 11 were selected for rotation. These 11 accounted for 85% of the variance in the items.

Later we quantitatively assessed the statistical significance of the unrotated factors by finding the correlation between the loadings of items for adjacent wavelengths (with each item included in only one pair). For factors that are continuous functions of wavelength the loadings on neighboring items should in general be similar, while pairs of randomly varying loadings will be uncorrelated. Of the 20 factors, 12 yielded significant correlations (the first 10, the 12th, and the 18th), but the magnitude of the loadings for the 18th was low. Thus only the first 10 to 12 are likely to be of theoretical interest, in agreement with our initial estimate.

After the Varimax rotation the first six factors were immediately physiologically interpretable; together these accounted for 70% of the total variance.[11] However, if only the most reliable measurements were considered (such as the amounts of the 526- and 645-nm primaries required to match yellow test lights), then the amount of variance explained rose to over 90%. The communalities based on these six factors are shown in Table 1 for each of the items.

The factor accounting for the most variance (16.9%) is shown in Fig. 1, where the observed loadings (unconnected symbols) are plotted as a function of the test light wavelength for each of the three primaries. (The abscissa is linear with wave number, reflecting the equal-wave-number intervals sampled by Stiles and Burch.) Note that the signs of the loadings change at the primary wavelengths (dashed vertical lines). This generally happens because at these wavelengths two of the primaries change sign as they are transferred from one half of the matching field to the other in order to permit the match. In this particular case it also reflects the change in whether the test light or the primary is more strongly absorbed by the pigment as the primary wavelength is crossed.

This factor was qualitatively identified as representing individual differences in macular-pigment density. Its clear importance as a source of variability in these data is somewhat surprising given the fact that the field size was 10° and subjects were instructed to ignore the central macular region defined by Maxwell's spot when matching the lights. Consistent with the absorption spectrum of the macular pigment, which peaks near 460 nm and is negligible at long wavelengths, the loadings for the red and green primary items for this factor are highest in the blue-green region of the spectrum and fall to near zero for yellow or red test lights. The exact relationship between variations in the pigment density and the resultant loadings is quite complicated. However, in broad outline it can be understood by noting that neither the 526- and 645-nm primaries nor the longer test wavelengths are absorbed to a great degree by the macular pigment, so that variations in the density cannot contribute much to the variability in the amounts of these primaries in matches to the longer-wavelength test lights. In contrast, at shorter wavelengths the pigment will selectively absorb the test lights, decreasing the logs of their effective radiances at the cone level (relative to the 526- and 645-nm primaries) in direct proportion to the amount of

Table 1. Communalities (Proportion of Variance Accounted for) Based on the 6 Identified Factors for the 97 Items Included in the Initial Factor Analysis[a]

Test Light				
Wave Number (cm^{-1})	Wavelength (nm)	Primaries R (645.2 nm)	G (526.3 nm)	B (444.4 nm)
14000	714.3	0.3958*	0.6519*	0.1531*
14500	689.7	0.5741	0.7909	0.2891
15000	666.7	0.3224*	0.6886*	0.2299*
16000	625.0	0.4880*	0.7742*	0.3898*
16250	615.4	0.8638	0.9419	0.6013*
16500	606.1	0.9181	0.9382	0.7970
16750	597.0	0.8849	0.9016	0.7372*
17000	588.2	0.9414	0.9466	0.8203
17250	579.7	0.8403	0.8600	0.8179*
17500	571.4	0.9203	0.3950	0.7276
17750	563.4	0.7957	0.7168	0.7529*
18000	555.6	0.9009	0.2418	0.8280
18250	547.9	0.7678	0.2769	0.7580*
18500	540.5	0.7656	0.1164	0.7374*
18750	533.3	0.7050*	0.3549*	0.7166*
19250	519.5	0.5725*	0.4449*	0.7184*
19500	512.8	0.7478	0.6275	0.7180
19750	506.3	0.8329	0.8508	0.8566
20000	500.0	0.7523	0.9070	0.8507
20250	493.8	0.8992	0.9125	0.8458
20500	487.8	0.8838	0.8968	0.9009
20750	481.9	0.9063	0.9380	0.9234
21000	476.2	0.8860	0.9382	0.7333
21250	470.6	0.8797	0.8306	0.5161
21500	465.1	0.8058	0.7884	0.4369
22000	454.5	0.4773*	0.4729*	0.3137*
23000	434.8	0.3945*	0.1950*	0.1988*
23500	425.5	0.5259	0.2458	0.6044
24000	416.7	0.7687	0.6601	0.6525
24500	408.2	0.9203*	0.8150*	0.9474*
25000	400.0	0.9592	0.9230	0.9640
25500	392.2	0.9155*	0.8845*	0.9504*
Age		0.5545		

[a] Asterisks indicate items that were deleted from the *a priori* rotation analysis to make room for the theoretically zero-loading items.

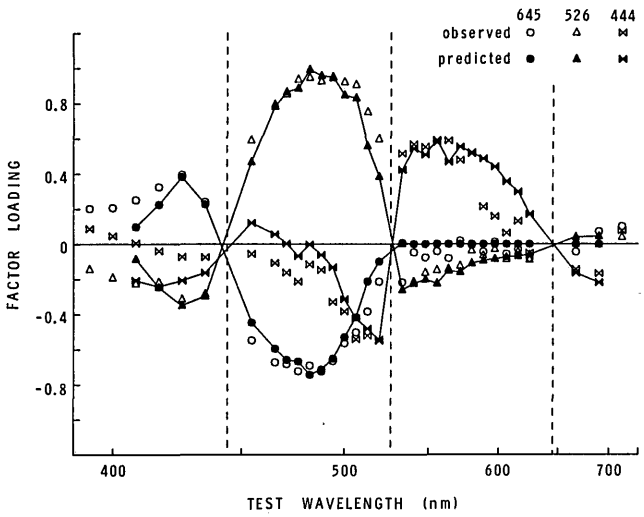


Fig. 1. Factor loadings for a factor corresponding to variations in macular-pigment density. Observed loadings (unconnected, open symbols) are plotted for each of the three primaries as a function of the matched test wavelength. Connected, filled symbols show the predicted loadings for macular density variations, with a standard deviation of 0.12 at 460 nm.

pigment present. Thus in this region differences in the density are a major source of the observed differences among subjects. The 444-nm primary behaves differently. In the blue-green region the loadings are near zero because both the test light and this primary are absorbed to a similar extent—a difference in density will not produce a relative difference in their radiances. It is only at longer test wavelengths, where the 444-nm primary is now selectively attenuated, that the influences of the pigment variations on this primary show up. At still longer wavelengths and at the shorter end of the spectrum, other factors become much more important as sources of variability in the blue primary (rod intrusion and lens pigment, respectively; see below) so that the proportion of variance explained by macular-pigment differences again diminishes. This final point illustrates that, while the derived factors are orthogonal (as a requirement of the Varimax rotation), it cannot be considered in isolation.

The polarity of the loadings is essentially arbitrary; rotating the relevant axes 180° changes their sign but not their magnitude. In this case the polarity is such that a lower density corresponds to a higher factor score. For example, an individual with a higher score on this factor would require more of the 526-nm primary and more negative amounts of the 645-nm primary (added to the test) to match the blue-green tests, as indicated by the positive and negative loadings, respectively, for these primaries at these wavelengths, and this results because a lower density makes the individual more sensitive to these tests.

The identification of this factor was confirmed by comparing the observed loadings with quantitative predictions of the effects of macular-pigment variations. As we noted above, the square of the factor loading for a given item equals the variance due to the factor divided by the total variance for the item. The loading itself is therefore the square root of this ratio, or the ratio of the respective standard deviations. Thus, if the effects of a known factor i on the measurements are theoretically determinable (and the variability among individuals in factor i is known), then the predicted loading of factor i on measurement j is

$$\hat{a}_{ji} = \sigma_{ji}/\sigma_j,$$

where σ_j is the observed standard deviation of item j and σ_{ji} is what the standard deviation of item j would be if all the variations in it were due to the ith factor. (An alternative would be to define σ_j as the square root of the sum of the variances due to each of the predicted factors. We rejected this alternative because it would assume that all the factors influencing the matches were known and included.) Now, for a given test wavelength λ, σ_{ji} is equal to the change in primary radiance, p_λ, per unit change in x_{ik} (a random variable representing the value of factor i for some individual k—for example, the macular-pigment density), multiplied by σ_{x_i}, the standard deviation of x_{ik} for the set of observers. Thus

$$\hat{a}_{ji} = \sigma_{x_i}(\partial p_\lambda/\partial x_i)/\sigma_j.$$

Finally, we chose to base all our predictions on variations in the logs of the primary radiances rather than on the linear radiances, so that our loadings were calculated from

$$\hat{a}_{ji} = \sigma_{x_i}(\partial \log|p_\lambda|/\partial x_i)/\sigma_{\log|j|}, \quad (1)$$

with the sign of the primaries restored after the logs of the absolute values were taken. The reason for using the logs of the radiances was that this allowed for predictions that were unaffected by assumptions about the particular spectra of other prereceptoral filters or the photopigments. These assumptions would affect the mean predicted primary radiances and their absolute variances but not, to a first approximation, the variance of their logs. Consider, for example, the effect of different lens-pigment spectra on the color matches calculated for a high or low density of the macular pigment. Any arbitrary change in the lens absorption will alter the pair of predicted primary radiances (for the high and low macular density), and also the difference between them, by the same multiplicative factor. However, the logs of the primaries will be changed by the same added constant, which cancels out when their difference is taken (and thus cancels from the calculated variance). Thus when the logarithms of the primary radiances are adopted as the observed variables, the predicted loadings are comparatively independent of any fixed but unknown scaling effects of other factors on the matching radiances.

In the case of the macular pigment, we chose to represent the value of the factor, x_{ik}, by the pigment density possessed by an individual at 460 nm, which is close to the wavelength of maximum absorption. An increase in x_{ik} changes the required log primary radiance $\log|p_\lambda|$ both by attenuating the test light and (in the opposite direction) by attenuating the primary stimulus itself. The net effect on the log matching intensity is proportional to the difference in the macular-pigment density at the test and primary wavelengths (relative to its density at 460 nm). Thus

$$\partial \log|p_\lambda|/\partial x_i = D_\lambda/D_{460} - D_p/D_{460}.$$

The density values were obtained from the macular-pigment absorption spectrum tabulated by Wyszecki and Stiles.[12]

This equation, when divided by the observed standard deviations in the log primary radiances, completely specifies all the predicted factor loadings for a factor representing individual variation in macular-pigment density, leaving as a free parameter only a multiplicative constant corresponding to σ_{x_i}, the standard deviation of the macular density among the observers. The IMSL subroutine ZXSSQ[13] was used to vary σ_{x_i} to obtain a least-squares fit of the predicted loadings to the observed loadings. This provided an estimate of 0.12 for the standard deviation of the macular-pigment density at 460 nm among the observers of Stiles and Burch. (The loadings do not provide information about the mean pigment density.)

The resultant predicted loadings are shown in Fig. 1 by the connected symbols and are clearly in close agreement with the observed loadings. This not only confirms the identity of the factor but also indicates that, at least in this case, the Varimax criterion was in fact sufficient to converge on the appropriate rotation without requiring any knowledge about the actual nature of the factors.

The second prereceptoral factor that we identified corresponded to individual differences in the density of the lens pigment and accounted for 12.1% of the total variance. In this case the loadings were characterized by high values for all three primaries at the shortest test wavelengths, falling to near zero at longer wavelengths (Fig. 2). To a first approximation this parallels the absorption spectrum of the lens pigment, as one would again predict from considerations of the effects of variations in its density. However, quantitative predictions based on the Wyszecki–Stiles[12] values for the lens-absorption spectrum yielded moderate loadings on the amounts of the blue and green primaries even at relatively long wavelengths, in part because of the low but not negligible density of the lens pigment for these primaries. For example, when the original predictions (not shown in the figure) were scaled to coincide with the observed loadings at short wavelengths, the predicted loadings for the blue and green primaries matched to yellow tests were roughly 0.4,

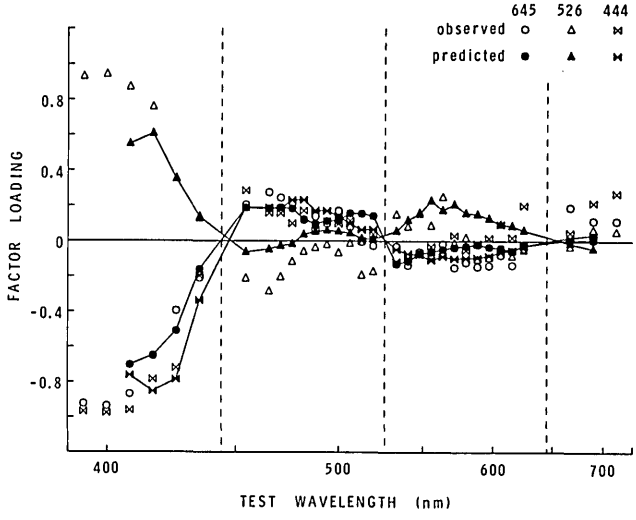


Fig. 2. Loadings for a factor identified as variation in the lens-pigment density. In this case the predicted loadings (connected, filled symbols) were first rotated with other predicted factors, using the Varimax rotation. The best-fitting standard deviation of the lens density is 0.18 at 400 nm.

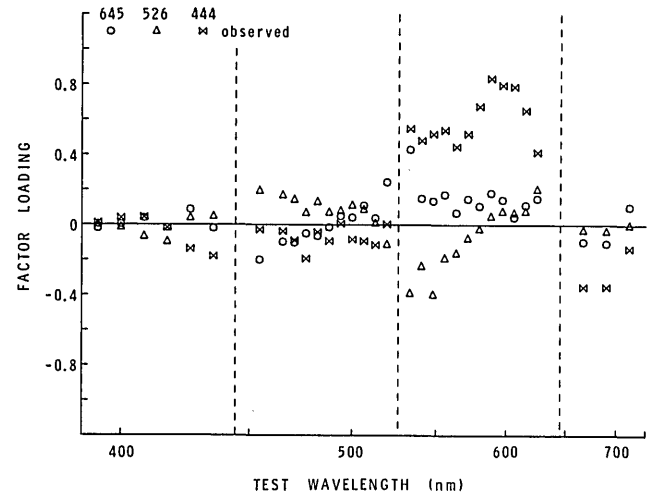


Fig. 3. Observed loadings for a factor corresponding to variation in the degree of rod intrusion, tentatively identified by the high loadings on the 444-nm primary matched to long test wavelengths. No predictions were calculated for this factor.

compared with less than 0.2 for the observed values. This pattern of loadings, which changes only gradually across much of the spectrum, was probably discouraged by the Varimax solution because of the low variance in the magnitudes of the loadings. To find out whether the rotation might in this case have been inappropriate, we performed a Varimax rotation of the entire set of *predicted* factor loadings. This did in fact change the lens-pigment predictions so as to bring them into close agreement with the set of observed loadings (as shown by the connected symbols in Fig. 2). The changes in the rotated loadings for the lens-pigment predictions obviously meant that other predicted factors had absorbed some of the variance that was due to lens variations. Nevertheless, except where noted below, the predictions for these other factors were not substantially affected by the rotation. The standard deviation of the lens-pigment density was estimated by a least-squares fit to be 0.18 at the reference wavelength of 400 nm (at which the mean lens density is 1.2 for the Wyszecki–Stiles tabulated values[12]).

In Fig. 3 the observed loadings have been plotted for a third factor, which accounts for 6.4% of the variance. It loads mainly on the amounts of the blue primary required to match long-wavelength tests, and for that reason we believe that it corresponds to the degree of rod intrusion. In color matches with large field sizes and intensities that are not extremely high, the rods may contribute differentially to the matches at longer wavelengths.[14] To a first approximation this results in a desaturation of the colors, requiring a compensating adjustment in the radiance of the blue primary. Differences in the extent of rod intrusion should thus account for much of the variability in the 444-nm primary at the longer wavelengths. Because of the large number of assumptions that it would require, we did not try to calculate predicted loadings for this factor.

The final three factors identified by our initial analysis appeared to represent variations in the long-, medium-, and short-wavelength-sensitive photopigments (hereafter referred to as L, M, and S, respectively) and accounted for 15.9, 10.5, and 8.5% of the total variance, respectively. One

possibility that we investigated was that these variations corresponded to variations in the density of the photopigments. An increase in the density alters the absorption spectrum (and hence changes the color matches) because it produces a relatively large increase in sensitivity in the tails of the spectra, with little gain at the peaks (because at the peaks the probability of photon capture is already high).[15] To examine whether this was the basis for the observed factors, we calculated predicted loadings for density variations using the photopigment spectral sensitivity estimates of Smith *et al.*[16] For a photopigment of sensitivity s_λ (normalized to a peak of 1.0) and initial density d_1, the normalized sensitivity at density d_2 can be calculated[15] from

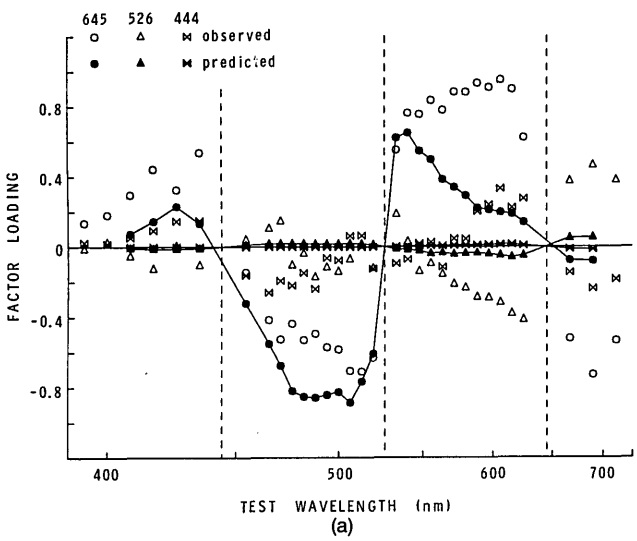
$$s_{d_{2\lambda}} = (1 - 10^{x})/(1 - 10^{-d_2}),$$

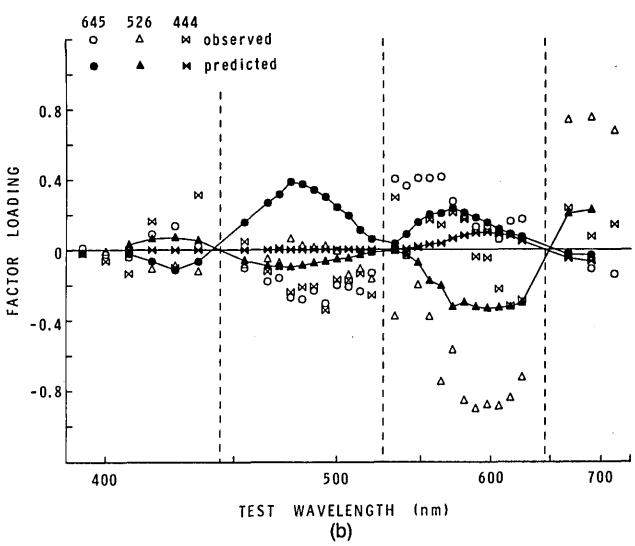
where

$$x = (d_2/d_1)\{\log[1 - (1 - 10^{-d_1})(s_{d_{1\lambda}})]\}.$$

Starting with the sensitivities of Smith *et al.*, new sensitivities were calculated for a ±0.05 change in the density of one of the pigments, and, by using these, two sets of color matches were derived. The difference in the logs of the matching primary radiances at the high and low densities, divided by the difference in density, provided an estimate of the derivative in Eq. (1). Finally, ZXSSQ was again used to find the standard deviation of density that yielded the best fit of the predicted to the observed loadings.

In Fig. 4(a) the predictions for density changes in the L pigment (connected symbols) are compared with the appropriate set of observed loadings, and it is obvious that the fit is poor owing to the negligible influence on the matches predicted for long test wavelengths. The fit of individual density variations in the M photopigment (to the loadings for the factor best described by that hypothesis) is still worse, as shown in Fig. 4(b). In this case the predictions fail largely because of the incorrect polarity predicted for the red primary matched to blue-green tests. The predictions for the S pigment turn out to be reasonably close to the corresponding S factor [Fig. 4(c)], but another interpretation of this factor is possible (see below). We also constructed a single theoretical factor based on a common density change in all three pigments, which might be more likely if density differences were due to individual differences in the lengths[17] or orientation[18] of all the cone outer segments. Yet here again the fit of this factor (shown in Fig. 7 below) to any of the three sets





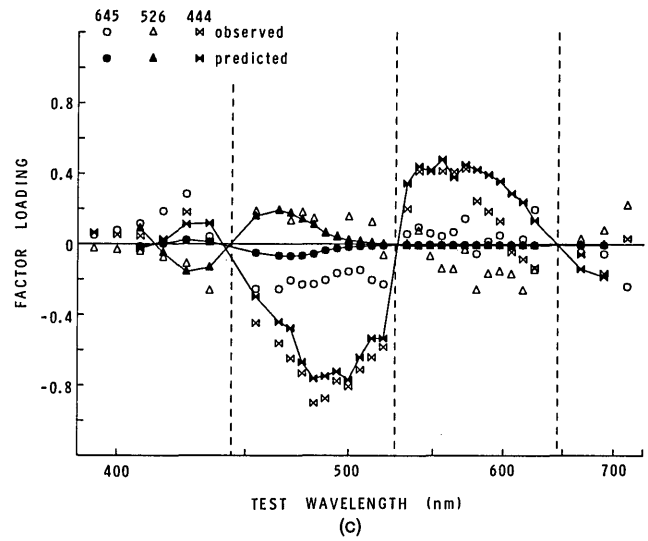


Fig. 4. (a) Unconnected, open symbols: observed loadings for a factor thought to reflect variations in the L cones. Connected, filled symbols: predicted loadings for density variations in the L pigment. The predictions are substantially different from the observed loadings at long test wavelengths, where they fail to predict the large observed influence on the matches. (b) Observed loadings for the M cone factor (unconnected, open symbols) compared with predictions for density variations in the M pigment. Predictions for the red primary matched to blue-green tests have the wrong polarity. (c) Observed loadings for the S cone factor (unconnected, open symbols) compared with predictions for density variations in the S pigment (connected, filled symbols).

of observed loadings was poor. Thus, with the possible exception of the S-pigment factor (further discussed and rejected below), none of these factors appears to represent individual differences in photopigment density.

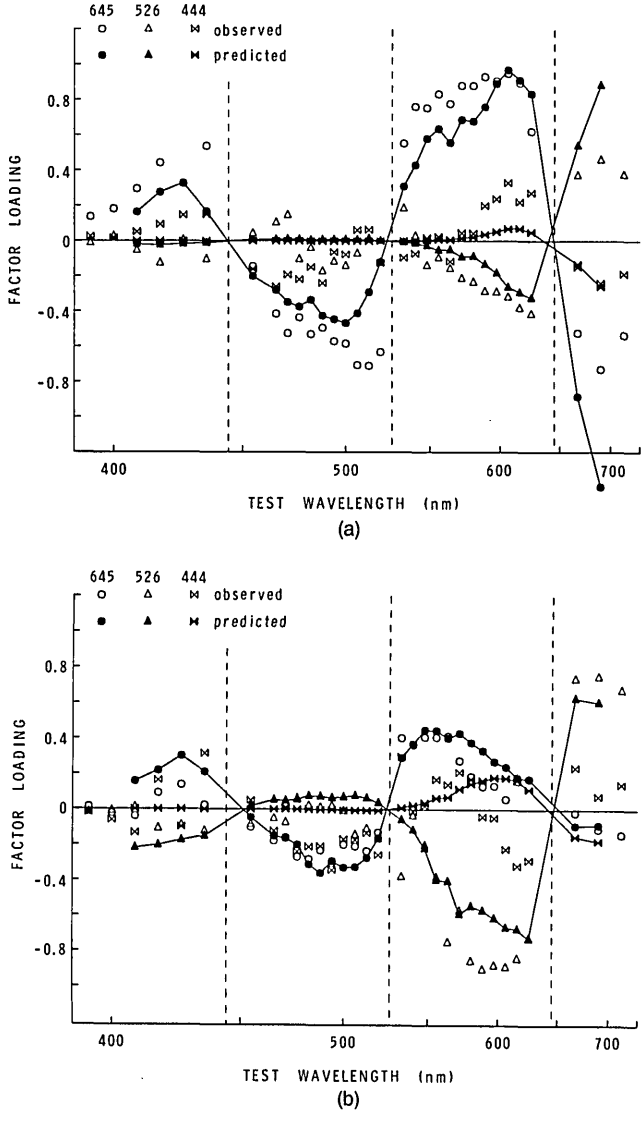
Instead, the observed factor loadings for all three factors correspond closely to predictions based on the hypothesis that the pigment variation takes the form of slight shifts in the absorption spectrum from one individual to another. To test this hypothesis, predicted loadings were constructed by shifting the spectra ± 250 cm^{-1} along the wave-number axis without changing their shape and calculating the resultant changes in the color matches. As shown in Figs. 5(a), 5(b), and 5(c) for shifts in the L, M, and S pigments, the fit of these predictions to the observed loadings is sufficiently close to argue strongly for λ_{max} variations as the correct factor identity. The required standard deviation of the peak was estimated by least-squares fits to be 50.3 cm^{-1} for the L pigment, 31.9 for M, and 45.3 for S. (Earlier informal estimates[5] differed slightly from these values.) Predicted loadings for a common shift factor, in which all three absorption spectra were displaced together in the same direction, did not resemble any of the three observed factors.

In the case of the S pigment the predictions for λ_{max} and density variations are roughly similar, so that it is difficult to discriminate between these alternatives on the basis of the predicted loadings of Figs. 4(c) and 5(c). However, the identification in a later analysis of a factor resembling common density variations in all three cone types (see below) argues against a separate and independent S-pigment density variation as the basis of the present factor.

Orthogonality of Factors

As we noted above, one property of the Varimax rotation is that the factors produced are all orthogonal to one another. However, the actual physiological factors need not necessarily be independent. For example, if variation in the three photopigment factors reflected some common influence, such as a general change in receptor morphology, then one would expect the factors to be correlated. In that case the Varimax rotation and the associated factor loadings would be in error. Alternatively, if the λ_{max} variation were due to differences in the actual photopigments, for example, because the genes encoding the pigments[19] varied slightly, then the corresponding factors would be expected to vary independently.

To examine whether the factors were truly orthogonal, we



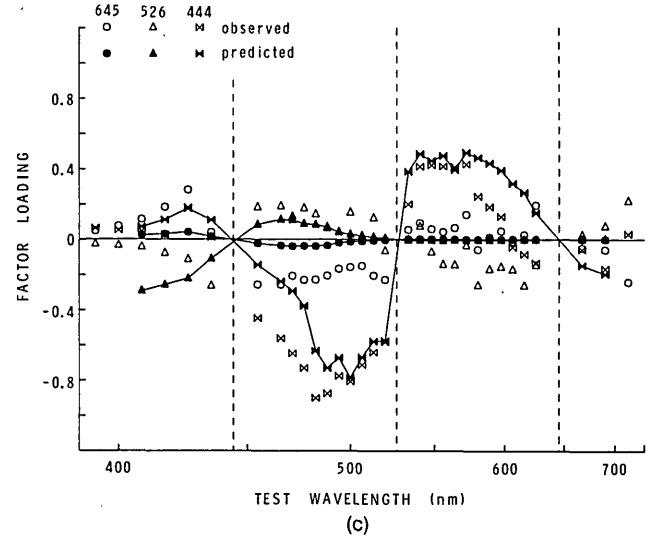


Fig. 5. (a) Loadings for the L cone factor (unconnected, open symbols), this time compared with predictions for variation in the spectral position of the L pigment (connected, filled symbols). The predictions are reasonably close to the observed loadings, suggesting that λ_{max} variations is the correct factor identity. The standard deviation in λ_{max} for these predictions is 50.3 cm^{-1}. (b) Loadings for the M cone factor compared with predictions for λ_{max} variations in the M pigment. The predictions (for a standard deviation of 31.9 cm^{-1}) closely follow the observed loadings. (c) Loadings for the S cone factor compared with predictions for λ_{max} variations in the S pigment (with a standard deviation of 45.3 cm^{-1}).

performed the analysis using the SPSS OBLIQUE rotation. In this rotation the parameter DELTA can be varied to control the degree of correlation among the factors. By setting it to 0.0 a solution with moderate interfactor correlations was actually encouraged. Nevertheless, the resulting correlations were all low and in most cases insignificant [with the exceptions of 0.38 for macular density and $\lambda_{max(S)}$ and of 0.28 for lens density and $\lambda_{max(S)}$]. We also used the same oblique rotation with the set of theoretical factors (macular density, lens density, and λ_{max} for each cone type). These were assumed to be orthogonal, yet an oblique rotation consistent with the predicted correlation matrix yielded interfactor correlations that were larger than those that we obtained using the empirical correlation matrix. (For the predicted factors the rms correlation was 0.33, compared with 0.16 for the empirical factors.) Thus the actual factors influencing the matches do indeed appear to vary independently. This implies, for instance, that there is no tendency for the densities of the lens and the macular pigments to covary (see also Fig. 4 of Ref. 2). Similarly, it suggests that the three photopigment factors do reflect variations that are restricted to each individual cone type.

Guided Factor Rotation

By relying on general statistical criteria to perform the factor rotation, we obviously assumed nothing about the specific mechanisms determining color matching or about what factors might be present. A better approach would be to take advantage of known properties of these mechanisms in order to help to guide the rotation. To do this we constructed a new set of variables by combining the original color-matching functions in ways that were designed to produce zero loadings on different types of factors. These zeros then served as targets to help to direct the rotation to a possibly more appropriate solution. Of course, if Varimax had arrived at the correct solution then the two procedures would have yielded similar results.

The derived variables were defined as follows. Eighteen items were created by taking, for each individual, the ratio of the radiance of two primaries at one test wavelength divided by the same ratio at a second wavelength. Such items are unaffected by differences in the densities of the prereceptoral pigments[20] and should therefore have zero loadings on the lens- and macular-pigment factors. For the rod intrusion factor, zero-loading items can be obtained simply by choosing the primary settings for test wavelengths at which the rod system is presumably driven to saturation. Avoiding the B primary settings where the rod intrusion is mainly reflected, we used 30 of the original R or G primary radiances matched to test wavelengths of 580 nm or less.[14]

To derive items that should be unaffected by variations in the photopigments, we relied on the fact that the color match can be regarded as being established independently within each of the cone types, when for each photopigment the absorptions due to the test light and the sum of the three primaries is the same.[21] For the Stiles–Burch data the match is thus defined by the following three equations:

$$\begin{aligned} s_{L_\lambda}Q_\lambda &= s_{L_{444}}Q_{444} + s_{L_{526}}Q_{526} + s_{L_{645}}Q_{645}, \\ s_{M_\lambda}Q_\lambda &= s_{M_{444}}Q_{444} + s_{M_{526}}Q_{526} + s_{M_{645}}Q_{645}, \\ s_{S_\lambda}Q_\lambda &= s_{S_{444}}Q_{444} + s_{S_{526}}Q_{526} + s_{S_{645}}Q_{645}, \end{aligned} \quad (2)$$

where s_{L_λ} is the sensitivity of the L pigment to test wavelength λ and Q_λ is the radiance of λ, etc. Now, suppose that two observers have the same S and M photopigments but differ in the sensitivity of their L pigments. Because s_{L_λ} differs for the two, the red, green, and blue primary radiances that they choose to match a particular test light will generally differ. However, for each of the pigments held in common, the chosen mixture of primaries must duplicate the excitation produced by the test light in order for a match to be achieved. Thus the different matches made by different observers will preserve equality of excitation for any pigments that the observers hold in common: increased amounts of one primary must be counterbalanced by decreased amounts of another, so that the excitation of the common pigment [determined as on the right-hand sides of Eqs. (2) by weighting each primary radiance by the sensitivity of the common pigment to that primary and summing the result for all three primaries] remains constant.[22] A factor that represents variation of any kind (for instance, in λ_{max}, density, or shape) in only one pigment should therefore have no influence (and zero loading) on such a measure of excitation of the unaffected pigments. (Of course, any two subjects may differ in any of or all their photopigment sensitivities. For the items isolating the contribution of one pigment to the match to have zero loadings on factors for the other pigments requires that the three photopigments vary independently, as our results suggest. See above.)

Eighteen of these new observed variables were calculated to represent excitation of either the L or the M photopigment as hypothetical common pigments; factors representing variations in the S pigment should have zero loadings on both of these sets. The items were constructed by weighting the three primaries matched to chosen test wavelengths by the L or M sensitivity of Smith *et al.*, adjusted to reflect the presumed effective optical density for 10° fields.[17] Along with the variables described above for prereceptoral factors, these were then entered into the standard analysis with Varimax rotation. Because of limitations in the total number of items permitted, the inclusion of these items necessitated deleting some of the original variables. Those removed have been indicated by asterisks in Table 1.

With the new Varimax factors as a starting point, the loadings were next modified by using an iterative procedure[23] in which we successively rotated the factor axes in pairs to try to line up the full matrix of observed loadings with the partial target matrix of theoretically defined zero loadings. The two factors of a pair were always rotated through a common angle to keep them orthogonal. Initially the angle was chosen to minimize the squared loadings on the particular set of targeted items (prereceptoral, photopigment, or rod intrusion) that were determined to be appropriate for each factor. However, this occasionally allowed one of the factors to be rotated into a plane in which *all* its loadings, and not just those for the theoretically zero-loading items, were near zero. To correct this, we therefore calculated the angle for each pair that minimized the ratio of the sum of the squared loadings on the targeted items to the total sum of the squared loadings for each factor. The procedure was repeated for each of the identified factors (paired with all other factors within a space of as many as 12 of the Varimax factors) until the rotation converged.

Despite the use of this entirely different rotation criterion,

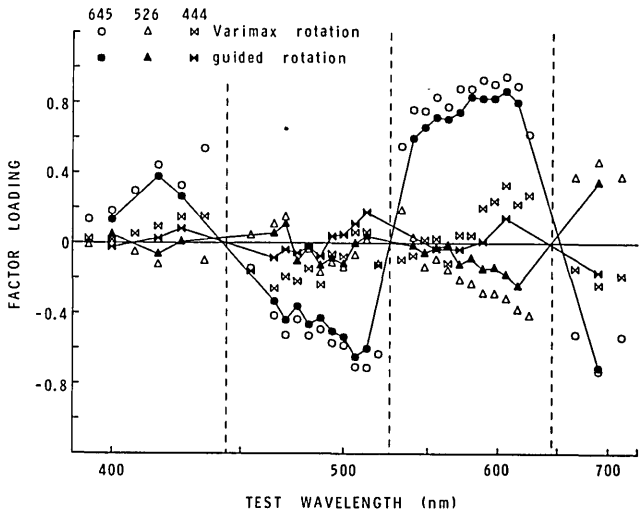


Fig. 6. Initial Varimax factor loadings for the L cone variations [unconnected, open symbols, same as Fig. 5(a)], compared with the revised loadings following the theoretically guided rotation of the factors (connected, filled symbols). There is remarkably little change in the loadings, indicating that the Varimax rotation for this factor is substantially correct and that the factor does represent variations in the L cones.

the loadings for most of the factors remained remarkably similar to the original Varimax estimates, indicating that the Varimax rotation was largely correct. This supports the identification of the different observed factors with variations in the prereceptoral pigments, the photopigments, and probably the participation of rods in the matches, all of which was based on the Varimax rotation. An example of the revised loadings is shown in Fig. 6, where the new $\lambda_{max(L)}$ loadings are plotted along with the Varimax loadings of Fig. 5(a). The two sets of loadings are similar throughout the spectrum. Thus the slight discrepancies between the observed pattern of loadings and the predicted loadings for $\lambda_{max(L)}$ variations remain [see Fig. 5(a)]. The source of these discrepancies is considered in the following paper.[4]

While this refined rotation using *a priori* assumptions about the nature of the operative factors largely confirmed our initial analysis, one significant modification in the results was the emergence of a factor resembling a common density variation in all the photopigments. The loadings for this factor are shown in Fig. 7, along with the predictions for common density variations as calculated above. The standard deviation of the density difference, which was held equal for each cone type, was estimated by a least-squares fit to be 0.045.

The loadings for this factor tend to vary little across much of the spectrum. As we have noted above, such a pattern is discouraged in the Varimax rotation, which probably obscured the presence of density differences in the initial analysis by allowing some of the variability that was due to them to be absorbed into the loadings for other factors. As a test of this, we examined whether our theoretical common density factor could survive a Varimax rotation in combination with the predicted factors for λ_{max} and prereceptoral differences. Several rotations were performed for a range of magnitudes of the initial density loadings (set by varying the standard deviation of density) to control the salience of the density factor. As expected, the effect of the rotation was to redistribute part of the variance due to density among the other factors, so that a pattern of loadings characteristic of density variations largely disappeared when the standard deviation in density was set to 0.03 or less. Again, however, the resulting change in the loadings for the other factors was slight. Thus, as the similarity between the initial and refined loadings of Fig. 6 suggests, the initial Varimax factors probably correspond closely to the actual physiological variations.

Predicting the Correlation Matrix

Thus far we have assessed the variance in the different hypothesized factors by fitting the loadings predicted by them to our set of observed factors. Yet, as we have repeatedly discussed, an incorrect rotation of the observed factors can introduce an error into the loadings by taking the variance that is really due to one factor and attributing it to another. If this happens then both the factor variance and how much the factor contributes to the total variance in the matches—as determined by the loadings—may be miscalculated.

One way around this problem is to choose the factor's standard deviations to fit directly the correlation matrix of the observed items rather than the particular set of loadings extracted from the matrix. In this case the question of whether the observed covariances are correctly partitioned never arises. (On the other hand, addressing the more practical problem of trying to *identify* the sources of variability by fitting the correlations would be problematic, as predictions for most potential factors would explain some of the variance.) The predicted correlation between two items, j and k, based on a set of n factors, is

$$\hat{r}_{jk} = \sum_{i=1}^{n} \hat{a}_{ji}\hat{a}_{ki},$$

where $\hat{a}_{ji}$ and $\hat{a}_{ki}$ are the predicted loadings of factor i on items j and k, respectively. The observed correlation then equals the predicted correlation plus a residual error term:

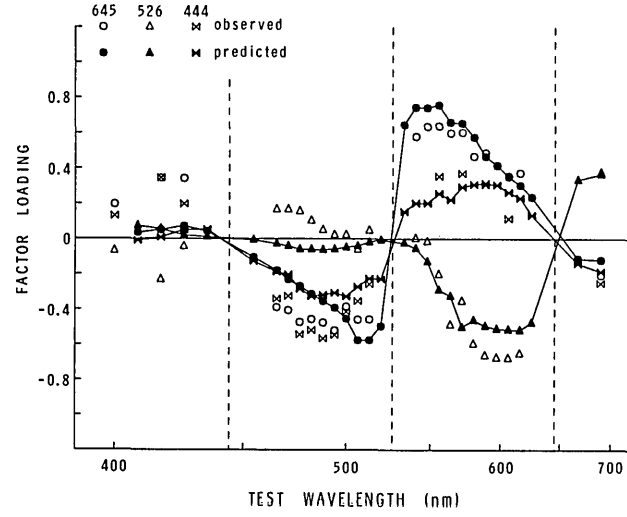


Fig. 7. Open, unconnected symbols: loadings for a factor that emerged from the theoretically guided rotation and qualitatively resembles predictions for a common density variation (standard deviation, 0.045) in all three pigments (connected, filled symbols).

Table 2. Estimated Interobserver Standard Deviations in Prereceptoral and Photopigment Density and in Photopigment λ_{max} (Wave Number)

	Macular	Lens	λ_L	λ_M	λ_S	Density
Fits to factor loadings						
10° Field color matches	0.12 (at 460 nm)	0.18 (at 400 nm)	50.3	31.9	45.3	0.045
2° Field color matches	0.18	0.15	47.2	25.0	41.3	—
Fits to correlation matrix						
10° Field color matches						
Independent λ_{max}	0.120	0.156	42.2	18.1	40.5	0.0458
Equal λ_{max}	0.121	0.153	32.6	32.6	32.6	0.0420
2° Field color matches						
Independent λ_{max}	0.184	0.231	48.0	25.1	24.1	0.0222

$$r_{jk} = \hat{r}_{jk} + e_{jk}.$$

For our set of factors we included predictions for macular and lens pigment differences, λ_{max} variation in the L, M, and S cones, and a change in photopigment density common to all three cones. Rod intrusion was not included because of the difficulties in making unambiguous predictions for this factor. The loadings for each factor were calculated as described above, using the standard deviation of each factor as a free parameter. ZXSSQ was then used to find the values of the standard deviations that would minimize the mean-squared error in the predictions.

The results of this procedure are shown in Table 2, where the best-fitting standard deviation for each of the identified factors is listed. For comparison, the standard deviations obtained by fitting the predicted loadings to the set of observed factor loadings are included in the first two rows. Also shown are the results of fitting either the correlation matrix or the factor pattern derived from an analysis of the Stiles–Burch 2° color matches,[3] which we discuss in the following subsection. Finally, for the fits to the correlation matrix for the 10° data two estimates of the variability of the predicted factors are shown, depending on whether the spectral peak factors were all constrained to have equal standard deviations or were allowed to vary independently.

Despite the fact that they did not show up among the initial set of factors, the present approach indicates that photopigment density variations do contribute significantly to the interobserver differences in the matches (as the guided rotation of the initial factors also suggested; see Fig. 7). The standard deviation in density required to fit the correlation matrix is roughly 0.045 for each pigment, a value close to the standard deviation suggested by the observed loadings of Fig. 7. Again, some of the variance due to density differences was probably absorbed into other factors in the Varimax solution in the way described above. Perhaps partly reflecting this, the standard deviations for the spectral peak factors fitted to the correlation matrix are somewhat lower than the estimates obtained by fitting to the observed factors. In particular, the standard deviation for the M pigment is less than 20 cm^{-1} when it is allowed to vary independently. The standard deviations for the S and L pigments are similar to each other but are almost double that of the M pigment. In a further analysis (also documented in Table 2) the three spectral peak factors were forced to have identical standard deviations, partly to test whether the value for the M pigment was due to competition with the density predictions for the same variance. However, this left the estimate for the density variations virtually unaffected, while producing a single intermediate value for the standard deviation of λ_{max}. The assumption of equal variability in λ_{max} across pigments is therefore not strongly supported.

The rms observed correlation in the original matrix was 0.39. After the variance due to the fitted predicted loadings was parceled out, the rms of the residuals, e_{jk}, dropped to 0.16. If the only relationships among the variables were due to the set of factors that we removed, then these residuals should be distributed in roughly the same way as correlations in samples from a population in which the correlation is zero. For 49 subjects the expected standard deviation of the zero correlation is $1/\sqrt{49-1}$, or 0.144, which is slightly less than the obtained value for the residuals of 0.16. This suggests that some real linkages between the variables still exist. As a final attempt at identifying potential sources of variability in the matches, we performed a factor analysis of the residual correlation matrix, following the procedure for our initial analysis above. One factor that we obviously expected to reappear was rod intrusion, and, in fact, a possible rod intrusion factor was again obtained, though it differed from the original by containing moderate loadings on long-test-wavelength items for the G primary as well as for the B primary. (A similar change in the rod intrusion factor was obtained in the guided rotation of the factors.)

Several other factors had loadings that varied systematically with wavelength and therefore potentially reflect some substrate of the color-matching process (if not systematic errors in some of our predicted loadings). However, we were unable to identify any of these. Among the possibilities that we considered was that of density variations in a filter with the absorption spectrum of the macular pigment, but screening only the L cones, which was postulated by Smith and Pokorny[24] to account for fits to Konig-type fundamentals and to explain individual variability in color matching.[16,25] The significant loadings predicted by this factor are restricted to the red primary items at mid to short test wavelengths. However, none of the observed factors from the residual matrix or from the initial set of 11 resembled this pattern. In any case, these unidentified factors account for only a small percentage of the variance in the Stiles–Burch data. Thus, whatever their possible theoretical significance, their contribution to the interobserver differences in these color matches is relatively minor.

Analysis of the Stiles–Burch 2° Color Matches

As we noted, Stiles and Burch[3] also measured color-matching functions for 2° fields. The data for the individual subjects were recently recovered and published by Tre-

zona.[26] We were interested in applying a similar analysis to the 2° data, both because these data provide an additional set of measures with which to examine individual differences and because they more closely reflect the color-matching properties of central vision. However, they have the drawback that the measures were made for only 10 observers, as opposed to 49 for the 10° fields.

For the 2° data, the correlation matrix was partitioned into nine components (one fewer than the number of observers), using principal-components analysis (which differs from factor analysis in that values of unity rather than estimated communalities are used in the diagonal elements).[27] As before, the factors were identified after Varimax rotation by comparing their loadings with the predicted loadings for a given theoretical factor.

The second row of Table 2 lists the best-fitting factor standard deviations that we obtained for each of the identified components. As before, we found factors representing macular- and lens-pigment variations and variations in the spectral peaks of the L, M, and S cones. There was no evidence of a rod intrusion factor, nor was there a clear photopigment density factor. A new far-red sensitivity factor, not clearly present in the analysis of the 10° matches, was associated with a need for greater energies of each of the three primaries matched to long-wavelength ($\lambda > 650$-nm) test lights. This could possibly be linked to individual variation in receptor morphology or refractive index, altering the waveguide properties of the receptors and thus the effective sensitivities of the cones for wavelengths that are large relative to the cone diameter. Variations in such a factor may be more apparent in the 2° data because of the smaller diameters of the foveal cones.

Probably because the observed loadings were based on so few subjects, the fit of our predictions was generally poorer than for the corresponding 10° field factors (particularly for the factors representing M and S cone variations). Nevertheless, the identity of the factors was clear (or at least suggestive for the M and S factors), and the estimated standard deviations were similar to the estimates based on the 10° data. The largest discrepancy, in the standard deviation of macular-pigment density, was in fact close to what we would have predicted from the effective mean density of the macular pigment for the two field sizes (see below).[28]

As with the 10° data, we also tried to estimate the factor standard deviations by fitting the correlation matrix for the 2° matches. The results are shown at the bottom of Table 2. These values are again in reasonable agreement with our other estimates, although if they are correct the lower variance in photopigment density would be surprising, given the greater length of the foveal cones.

DISCUSSION

To summarize, we have identified the following sources of interobserver variability in the Stiles–Burch 10° color matches: differences in the lens- and the macular-pigment densities, in the degree or form of rod intrusion, and in the density and spectral peak of the photopigments. All these factors have been known or previously postulated to underlie the observed variability in color matching. Numerous other physiological factors no doubt exist, but it is unlikely that the overall contribution of any of them is large, because the variance of the majority of items is largely accounted for by the set of identified factors. For the most reliable matches the communalities based on these factors are near 0.90, providing close to a complete description of the intersubject differences. Thus, for example, if one were to rank the subjects for these items (according to how much of the primary they would require to match the test) simply on the basis of their combined factor scores, then the predicted rank would correlate $\sqrt{0.90}$, or 0.95, with the observed rank.

For our initial set of factors we noted the percent of the total variance that was accounted for by each. However, it should be kept in mind that this number is dependent on which variables are included. For instance, if the item set contained more short-wavelength tests, then the proportion of variance due to the inert pigments would obviously increase. As we noted above, the Stiles–Burch data include twice as many wavelengths from the midspectral region as from either extreme. Recall, moreover, that the percent of the total variance, as commonly defined, reflects the sum of the *standardized* variances for each item. However, the matches to some wavelengths have much more absolute variability associated with them than others. The factor loadings contain no information about what the absolute variance is, indicating only how much of it is due to the factor. The actual variations in primary radiance that are attributable to the factor can be found by multiplying the loadings by the observed standard deviations. As an example, Table 3 shows the standard deviations (expressed as percentages of the mean settings) in the radiance of the primaries matched to 588 and 482 nm that would be created by the estimated variation in each of our factors from the initial Varimax rotation. (The complete list of observed standard deviations can be found in Ref. 2.)

Our results suggest a standard deviation of 0.12 for the macular-pigment density at 460 nm and from 0.15 to 0.18 for the lens pigment at 400 nm (depending on whether the correlation matrix or the factor loadings were predicted). These values are fairly comparable with previous estimates. For example, Bone and Sparrock[29] reported a standard deviation of 0.15 to 0.20 for the macular density of 49 subjects measured with 2° fields. For a 2° matching field the mean density at 460 nm is roughly 0.5.[12] However, Stiles and Wyszecki estimate that the effective density for 10° fields is roughly 0.6 times that for 2°.[28] Our estimate of 0.12 might therefore suggest a standard deviation of 0.20 for the smaller field size and thus is very close to the value of 0.18 we actually obtained for the 2° matches (see Table 2).

Van Norren and Vos[30] calculated presumed lens-density spectra from the scotopic sensitivity data of Crawford and found that 95% of the observers (between 17 and 30 years of age) fell within ±25% of the mean density. This is consistent with our estimate of ~0.15 as the standard deviation in density at 400 nm if we use the Wyszecki–Stiles[12] mean density of 1.2 at 400 nm; it suggests a slightly smaller range of ±20% if the van Norren–Vos higher estimate of 1.45 is used. Both of these mean values may underestimate the average value for the Stiles–Burch subjects. Pokorny *et al.*[31] have suggested that the Wyszecki–Stiles density function should be multiplied by 1.33 to adjust it for the average age of this sample, which would reduce our estimate of the percentage change accordingly. Note that the standard deviation of 0.15 is 12.5% of 1.20 (our lowest mean density

Table 3. Standard Deviations (σ) in the Color-Matching Functions That Would Be Generated by the Estimated Variation in the Factors Identified in the Initial Analysis[a]

Test Light	Primary	Total Observed	Macular	Lens	λ_L	λ_M	λ_S	Rod Intrusion
588 nm	R	9.1	0.41	1.3	8.5	1.2	0.15	1.7
	G	6.5	0.43	0.68	1.9	5.9	1.1	0.31
	B	22	4.7	0.31	4.4	0.86	4.1	18
482 nm	R	15	10.4	1.8	7.8	4.2	3.4	0.96
	G	10	9.5	0.61	0.39	0.23	1.4	1.3
	B	9.4	1.1	1.6	1.4	2.0	8.5	0.40

[a] The values are expressed as percentages of the mean matching intensities.

estimate), while that of 0.12 for the macular density is 40% of the presumed 10° mean of 0.3. Thus the relative variability in the macular-pigment density is probably at least three times as large.

There is considerable evidence that the density of the lens pigment increases with age.[32] However, subject age did not significantly correlate with our lens-density factor ($r = 0.24$) and was only weakly correlated with other factors. [The highest correlations were 0.44, 0.40, and 0.34 for $\lambda_{max(S)}$, $\lambda_{max(M)}$, and macular pigment, respectively.] Stiles and Burch similarly reported only a weak relationship between lens density and age for these data.[2] This is perhaps because there were few very young or old observers in this sample (mean age, 31.8; standard deviation, 10.3), and differences within age groups were large. Moreover, we have modeled the individual differences in lens pigment by uniformly scaling the density of the Wyszecki–Stiles absorption spectrum. However, it has been suggested that this function might be composed of two separate components, only one of which shows significant changes over the age range of this sample.[31]

The fitted predictions for optical density variations in the photopigments indicate a standard deviation of roughly 0.045. Only a single density factor was suggested, reflecting a correlated variation in the pigments. In contrast, three λ_{max} factors were obtained, which represent independent variations in the three cone types. A somewhat related result with respect to density was found by Smith *et al.*[16] They evaluated the range over which the density and the spectral peak of the L and M cones could vary and still produce color matches that were within the range of individual differences of the Stiles–Burch 2° matches.[3] While the acceptable λ_{max} values for one cone type depended little on the other, the fit to the observed variations required that the densities of the two pigments covary. A common density variation could possibly reflect a general variation in the size of all the cone outer segments. The standard deviation of 0.045 would correspond to roughly a 3-μm change in the outer segment lengths.[33]

In Table 4 we have reproduced our estimates of the standard deviations in each of the spectral peak factors, this time expressed in wavelengths (at λ_{max}). Also shown are several other psychophysical and physiological measures of the variability in λ_{max}. Alpern[34] has calculated presumed λ_{max} distributions for a population of protanopes and deuteranopes, based on the anomolascope settings made by Alpern and Wake[35] and Bastian.[36] From these we have estimated that the standard deviation in the L and M peaks is roughly 1.4 and 2.0 nm, respectively. Using flicker photometry on colored adapting backgrounds to try to isolate the cone sensitivities, Eisner and MacLeod[37] found that their group of seven normals clustered about peak L cone sensitivities 85 cm^{-1} (2.7 nm) apart, suggesting a standard deviation of about half of this amount. Finally, Smith *et al.*[16] calculated that deviations covering a range of 300 cm^{-1} (9.3 nm) in the L cones and 200 cm^{-1} (5.7 nm) in the M cones yielded predicted matches that were within the extremes of the Stiles–Burch 2° matches (though they also argued that variation in density with no variation in λ_{max} provided an equally good account of the data that they considered).

Baylor *et al.*[38] have measured the spectral sensitivities of the cones by directly recording the outer segment photocurrents. The standard deviations that they obtained for their sample of L, M, and S cones were 1.0, 1.3, and 1.4 nm, respectively. Another approach has been to measure the light absorption of individual cones by using microspectrophotometry (MSP). The MSP data shown in Table 4 are the standard deviations in λ_{max} found from estimates of the sensitivities of both human[39] and Old World Monkey cones.[40,41] These values, which range from 2 to 5 nm, tend to be larger than the corresponding psychophysical estimates as well as those based on recording the cone responses. Of course, the psychophysical estimates reflect the differences

Table 4. Variability (Standard Deviation in Nanometers) in the Estimated Spectral Peaks of the Photopigments Reported by Several Psychophysical and Physiological Studies

Study	Variability (nm) L	M	S
Psychophysical Estimates			
This study	1.5	0.9	0.8
Alpern[34]	1.4	2.0	
Eisner and MacLeod[37]	1.3		
Smith *et al.*[16]	9.3[a]	5.5[a]	
Physiological estimates			
Baylor *et al.*[38] (*Macaca fascicularis* data)	1.0	1.3	1.4
Dartnall *et al.* (human data)[39]	3.6	3.5	5.2
Bowmaker *et al.* (*M. mulatta* data)[40]	4.8	4.8	
MacNichol *et al.* (*M. fascicularis* data)[41]	2	2	3.5

[a] Measures of the full range.

between large populations of cones in different individuals, while the MSP is based on differences in single receptors. In any case, the standard deviations that we have obtained suggest an upper limit on the expected interobserver variance. The standard deviations in peak sensitivity could conceivably be smaller if the λ_{max} factors included some variance due to other factors, but they are unlikely to be significantly larger, as this would exceed the observed variability in the matches. On the other hand, large variability of the pigments within a single individual is inconsistent with the known additivity of color matches, because the matches should then break down as different subpopulations are selectively adapted. For example, Nagy *et al.*[42] have calculated that a standard deviation of greater than 1.6 nm is inconsistent with the constant Rayleigh matches that they found for different adapting backgrounds.

Our conclusion that there is less jitter in the λ_{max} for the M cones than for the L cones is consistent with the results of Neitz and Jacobs[43] but differs from the findings of Alpern.[34] Nathans *et al.*[19] report that many individuals have multiple copies of the M-pigment gene. If differences in the genes were the basis of the variation, then simultaneous expression of these copies would reduce variability in the effective (average) λ_{max} of the M cones. But deuteranomalous observers who possess apparently normal genes among their copies of the M-pigment gene (for example, author DIAM)[19] seem to lack the normal M pigment, and this makes the hypothesis of simultaneous expression unlikely.

Several studies have suggested that the variations in λ_{max} are not normally distributed. With regard to human pigments, for example, Dartnall *et al.*[39] noted that the MSP results for the λ_{max} of both the M and L cones appear to form bimodal distributions (though only the L distribution is statistically nonnormal). As we mentioned above, Eisner and MacLeod[37] found that the L cone peaks of their observers clustered into two groups, and recently Neitz and Jacobs[43] argued from an analysis of Rayleigh matches that the L cone spectra in males are bimodally distributed (with a separation of roughly 3 nm in the peaks). The estimated distributions of λ_{max} for the Stiles–Burch observers are shown in Fig. 8, with the individual values for the 34 male and 15 female subjects plotted as downward or upward displaced lines. These distributions represent the factor scores for each of the λ_{max} factors obtained from a principal-components analysis of the item set with the theoretically zero-loading items (which again yielded factors similar to the initial Varimax factors.) The scores have a mean of 0 and a standard deviation of 1. An individual's score is thus a relative measure of how many standard deviations he falls from the mean. The actual presumed distribution of λ_{max} can be obtained by multiplying the scores by the estimates of the λ_{max} standard deviations for each factor (in Table 2).

Following Dartnall *et al.*,[39] we applied the Shapiro–Wilk[44] test to determine whether the factor scores for the λ_{max} and other factors are normally distributed, analyzing all scores together as well as the male and female scores separately. However, only the distribution of $\lambda_{max(M)}$ for males is nonnormal by this criterion ($W = 0.929$ for $n = 34$, $p < 0.05$), and in this case there is no obvious clustering of the scores. Thus the factor scores give no real indication of separate subpopulations of each pigment type. We also compared the relative spread in the scores for male and female observers. Suppose

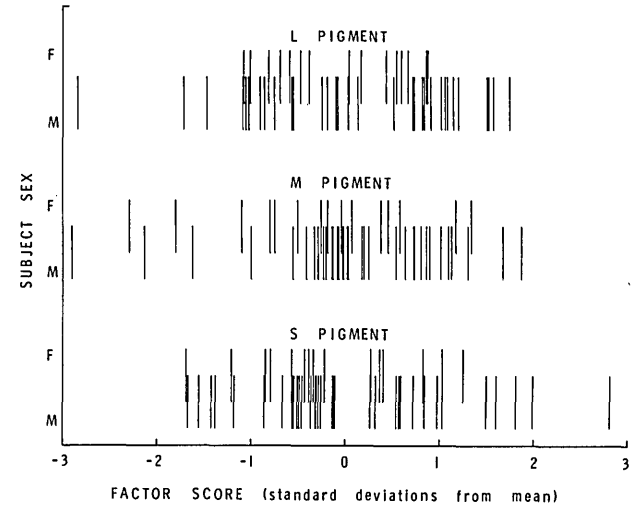


Fig. 8. Individual factor scores for the factors representing λ_{max} variations in the L, M, and S pigments. Each score indicates how many standard deviations the individual's λ_{max} fall from the mean for the group, with positive deviations corresponding to longer-wavelength peaks. For each factor the scores for females have been displaced upward, the male scores downward.

that the variations in spectral peak were due to differences in the protein sequences for the genes coding the pigments. Because the L and M genes are X linked, females have two alleles for each, which are independently inherited and independently inactivated in each cone.[45] Their effective λ_{max} should therefore be roughly the average of the two, and the female variance in λ_{max} should be half that of the males. In contrast, the male and female variances for the autosomal S gene should be similar. The actual ratio of male/female variance in the factor scores was 2.57 for L, 1.03 for M, and 1.63 for S. While these appear to differ from the predictions for a simple genetic basis for the λ_{max} variation, they are in fact not inconsistent with it, because the critical F ratio for evaluating them is itself well above 2 ($F = 2.41$, $df = 33, 14$).

Our results for the Stiles–Burch 2° color matches are largely similar to those that we obtained for the 10° matches. As before, we found three factors that appear to reflect independent variations in the λ_{max} of the three photopigments (though only the L pigment factor was clearly defined), and the estimated variability in λ_{max} was close to the values that we estimated from the 10° data. Smith *et al.*[16] and Estevez[46] have also suggested that λ_{max} variations might partly contribute to the individual differences in these data, though they could not distinguish these from possible photopigment density variations. None of the extracted factors that we found clearly corresponded to differences in photopigment density, and only a weak role of density differences was suggested by the fits to the correlation matrix. Thus for both sets of data λ_{max} variations appear to be the more salient determinant of individual differences in the matches.

ACKNOWLEDGMENTS

We thank F. J. J. Clarke of the National Physical Laboratory for providing us with the individual color-matching data from the study of Stiles and Burch. This study was supported by National Institutes of Health grant EY01711.

* Present address, Department of Experiment Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK.

Address reprint requests to D. I. A. MacLeod.

REFERENCES AND NOTES

1. W. D. Wright, *Researches in Normal and Defective Colour Vision* (Kimpton, London, 1946).
2. W. S. Stiles and J. M. Burch, "N.P.L. colour matching investigation: final report (1958)," Opt. Acta **6,** 1–26 (1959).
3. W. S. Stiles and J. M. Burch, "Interim report to the Commission Internationale de l'Eclairage, Zurich, 1955, on the National Physical Laboratory's investigation of colour-matching," Opt. Acta **2,** 168–181 (1955).
4. D. I. A. MacLeod and M. A. Webster, "Direct psychophysical estimates of the cone-pigment absorption spectra," J. Opt. Soc. Am. A **5,** 1736–1743 (1988).
5. A preliminary account of this work appeared in D. I. A. MacLeod and M. A. Webster, "Factors influencing the color matches of normal observers," in *Colour Vision Physiology and Psychophysics,* J. D. Mollon and L. T. Sharpe, eds. (Academic, London, 1983) pp. 81–92.
6. Because of their greater familiarity, we have tried to use measures of wavelength whenever practical. The wave number is the reciprocal of the wavelength expressed in centimeters.
7. In the actual study a yellow primary at 588 nm was substituted for the green for test wavelengths longer than 588 nm. Further, subjects were run in two separate groups, which differed in the intensity levels used and the location of the blue primary (445.4 versus 470.6 nm). The correction factors supplied by Stiles and Burch for slit width differences, etc. between the two groups were applied to the data to make values for the two sets comparable.
8. General treatments of factor analysis are available in numerous texts. We relied mainly on H. H. Harman, *Modern Factor Analysis* (U. Chicago Press, Chicago, Ill., 1960).
9. N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbrenner, and D. H. Bent, *Statistical Package for the Social Sciences,* 2nd ed. (McGraw-Hill, New York, 1975).
10. Because the correlation matrix is so large and the communalities are high, the choice of values for the diagonal elements of the matrix is not critical for the results. A principal-components analysis in which values of 1 are assumed for these elements was also tried, and the results were not substantially different from those obtained by a true factor analysis with iteratively estimated communalities.
11. Our initial report of 75% was based on the first six factors before rotation.
12. G. Wyszecki and W. S. Stiles, *Color Science* (Wiley, New York, 1967).
13. *IMSL Library* (IMSL, Houston, Tex., 1984).
14. W. S. Stiles and G. Wyszecki, "Rod intrusion in large-field color matching," Acta Chromatica **2,** 155–163 (1973).
15. G. Wyszecki and W. S. Stiles, "High-level trichromatic color matching and the pigment-bleaching hypothesis," Vision Res. **20,** 23–37 (1980).
16. V. C. Smith, J. Pokorny, and S. J. Starr, "Variability of color mixture data—I. Interobserver variability in the unit coordinates," Vision Res. **16,** 1087–1094 (1976).
17. J. Pokorny and V. C. Smith, "Effect of field size on red–green color mixture equations," J. Opt. Soc. Am. **66,** 705–708 (1976).
18. G. S. Brindley, "The effects on colour vision of adaptation to very bright lights," J. Physiol. **122,** 332–350 (1953).
19. J. Nathans, D. Thomas, and D. S. Hogness, "Molecular genetics of human color vision: the genes encoding blue, green, and red pigments," Science **232,** 193–202 (1986).
20. J. von Kries, in *Helmholtz's Treatise on Physiological Optics* (Optical Society of America, 1924; Dover, New York, 1962), Vol. II, pp. 407–408.
21. J. C. Maxwell, "Experiments on colour, as perceived by the eye, with remarks on colour-blindness," Trans. R. Soc. Edinburgh **21,** 275–298 (1855); T. N. Cornsweet, *Visual Perception* (Academic, New York, 1970).
22. D. I. A. MacLeod and M. Hayhoe, "Three pigments in normal and anomalous color vision," J. Opt. Soc. Am. **64,** 92–96 (1974).
23. M. W. Browne, "Orthogonal rotation to a partially specified target," B. J. Math. Statist. Psychol. **25,** 115–120 (1972).
24. V. C. Smith and J. Pokorny, "Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm," Vision Res. **15,** 161–171 (1975).
25. J. Pokorny, V. C. Smith, and S. J. Starr, "Variability of color mixture data—II. The effect of viewing field size on the unit coordinates," Vision Res. **16,** 1095–1098 (1976).
26. P. W. Trezona, "Individual observer data for the Stiles–Burch 2° pilot investigation," J. Opt. Soc. Am. A **4,** 769–782 (1987).
27. For the 2° matches Stiles and Burch measured the full spectrum in 250-cm^{-1} steps from 392 to 727 nm, giving us 135 variables for the three primaries.
28. W. S. Stiles and G. Wyszecki, "Colour-matching data and the spectral absorption curves of visual pigments," Vision Res. **14,** 195–207 (1974).
29. R. A. Bone and J. M. B. Sparrock, "Comparison of macular pigment densities in human eyes," Vision Res. **11,** 1057–1064 (1971).
30. D. Van Norren and J. J. Vos, "Spectral transmission of the human ocular media," Vision Res. **14,** 1237–1244 (1974).
31. J. Pokorny, V. C. Smith, and M. Lutze, "Aging of the human lens," Appl. Opt. **26,** 1437–1440 (1987).
32. J. S. Werner and B. R. Wooten, "Age changes in ocular media density and consequences for colour vision," in *Colour Vision Deficiencies V,* G. Verriest, ed. (Higer, Bristol, UK, 1980) pp. 355–359.
33. R. W. Rodieck, *The Vertebrate Retina* (Freeman, San Francisco, Calif., 1973).
34. M. Alpern, "Variation in the visual pigments of human dichromats and normal human trichromats," in *Frontiers of Visual Science: Proceedings of the 1985 Symposium* (National Academy Press, Washington, D.C., 1987), pp. 169–193.
35. M. Alpern and T. Wake, "Cone pigments in human deutan colour vision defects," J. Physiol. **266,** 595–612 (1977).
36. B. L. Bastian, "Individual differences among the photopigments of protan observers," Ph.D. dissertation (University of Michigan, Ann Arbor, Mich., 1976).
37. A. Eisner and D. I. A. MacLeod, "Flicker photometric study of chromatic adaptation: selective suppression of cone inputs by colored backgrounds," J. Opt. Soc. Am. **71,** 705–718 (1981).
38. D. A. Baylor, B. J. Nunn, and J. L. Schnapf, "Spectral sensitivity of cones of the monkey, *Macaca fascicularis,*" J. Physiol. **390,** 145–160 (1987).
39. H. J. A. Dartnall, J. K. Bowmaker, and J. D. Mollon, "Microspectrophotometry of human photoreceptors," in *Colour Vision Physiology and Psychophysics,* J. D. Mollon and L. T. Sharpe, eds. (Academic, London, 1983), pp. 69–80.
40. J. K. Bowmaker, J. D. Mollon, and G. H. Jacobs, "Microspectrophotometric results for old and new world primates," in *Colour Vision Physiology and Psychophysics,* J. D. Mollon and L. T. Sharpe, eds. (Academic, London, 1983), pp. 57–68.
41. E. F. MacNichol, J. S. Levine, R. J. W. Mansfield, L. E. Lipetz, and B. A. Collins, "Microspectrophotometry of visual pigments in primate photoreceptors," in *Colour Vision Physiology and Psychophysics,* J. D. Mollon and L. T. Sharpe, eds. (Academic, London, 1983), pp. 13–38.
42. A. L. Nagy, D. I. A. MacLeod, N. E. Heyneman, and A. Eisner, "Four cone pigments in women heterozygous for color deficiency," J. Opt. Soc. Am. **71,** 719–722 (1981).
43. J. Neitz and G. H. Jacobs, "Polymorphism of the long-wavelength cone in normal human colour vision," Nature **323,** 623–625 (1986).
44. S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," Biometrica **52,** 591–611 (1965).
45. M. F. Lyon, "Gene action in the X-chromosome of the mouse (*mus musculus L.*)," Nature **190,** 372 (1961).
46. O. Estevez, "Interindividual receptor variability of normal colour observers: analysis of the 2-deg Stiles and Burch data," Perception **15,** 677–687 (1986).