Modeling trichromatic color appearance with only two spectrally distinct photopigments

A. Kimball Romney[1*] and Chuan-Chin Chiao[2]

[1]Institute for Mathematical Behavioral Sciences, University of California, Irvine, Irvine, CA 92697, USA.

[2]Department of Life Sciences, National Tsing Hua University, Hsinchu, 30013, Taiwan.

[*]To whom correspondence should be addressed; E-mail: akromney@uci.edu

Protanomaly is a common abnormality of color vision in which the long wavelength sensitive cone is lacking. Protanomalous observers are trichromatic with reduced ability to discriminate colors in the red-green part of the spectrum because the photopigments that mediate discrimination in this range are abnormally similar differing in one of two ways; first, by only a few nanometers in peak sensitivity, or second, differing only in optical density. We present a functional model that estimates the color appearance and spectral color boundaries for these two types of protanomaly. The model predicts rather similar appearance of Munsell color chips but radically different spectral color characteristics for the protanomalous observers compared to normal color observers.

Neitz *et al.*[1] provide a detailed investigation of two different types of protanomalous trichromat observers lacking a long wavelength sensitive cone. Unlike normal observers with peak sensitivities between long and medium wavelength cones of about 30 nm; the protanomalous subjects have two genes for photopigments of medium wavelength sensitivity cones with unusual characteristics. The first type (P3nm), represented by subjects 14 and 15, is a protanomalous trichromat with variants of the medium wavelength sensitive pigments separated by only three nanometers (530 nm and 533 nm). The second type (P0nm), represented by subject 20, is a protanomalous trichromat with variants of the medium wavelength sensitive pigments with identical peaks at 530 nm but differing in optical density. Neitz *et al*. reported that the protanomalous trichromats they studied were able to make limited color discriminations in the red-green region of the spectrum, but that this ability was greatly reduced compared to normal observers. Here we show how to estimate the color appearance of Munsell color chips and monochromatic spectral lights for the two types of protanomalous observers described by Neitz *et al.* based on a recently proposed model[2,3] for estimating color appearance.

Fig. 1 shows the distribution of cone sensitivities for a normal observer defined by Stockman and Sharpe[4] compared to the two types of protanomalous observers reported by Neitz *et al*. The protanomalous curves were estimated using the peak locations reported by Neitz *et al* and the alpha band template formula of Govardovskii *et al.*[5] The broader curve at 530 nm in Fig. 1C simulates a density difference by increasing the standard deviation of the template curve by taking the square root at each nm value.

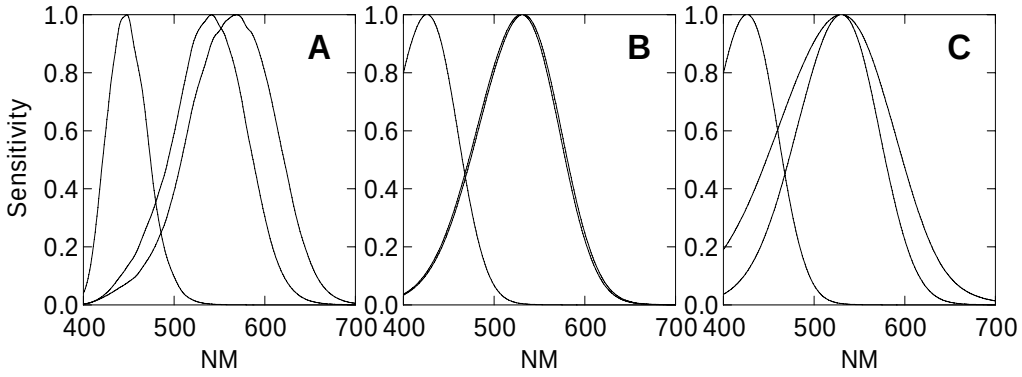

Fig. 1. Three examples of photoreceptor arrangements; (A) normal observer, (B) P3nm observer with peak sensitivities at 426 nm, 530 nm, and 533 nm, (C) P0nm observer with peak sensitivities at 426 nm and two peaks at 530 nm simulating optical density differences.

The three photoreceptor arrangements shown in Fig. 1 represent examples of individual differences among observers related to genetic differences in photopigments. The model may be used to estimate color appearance for observers with any genetic combination of photopigments. For example, in a previous paper[2] we modeled the three types of dichromats who completely lack one of the three photoreceptors, and accurately predicted matching spectral lights for a unilateral tritanope. The aim of this paper is to use the computational model to account for the seemingly unattainable trichromatic color vision in these protanomalous trichromat observers. Other theoretical accounts of anomalous trichromats in predicting Rayleigh matching have been reported, for example, variations among deuteranomalous trichromats are discussed by Shevell and He[6] and He and Shevell[7]. Variations also extend to the contribution of rods to color vision as in the study by Shevell *et al.*[8].

The model we use is described in detail by Romney and Chiao[2] with additional data smoothing corrections suggested by Romney[3]. We emphasize that even though the model is not intended to imply any physiological analogs, we do claim that computations that result in outcomes similar to those of the model are carried out by the human visual system. Here we present the essential components of the model with minimum replication of clarifying explanations. To obtain empirical predictions we need to focus on one or another of our specific human observers and a specific input data set of surface reflectance spectra. The receptor data on our selected observers shown in Fig. 1 are denoted as the matrix, $\mathbf{R}_{301\times3}$ in Eq. 1. The input data consist of the same sample developed by Romney[3] of 640 smoothed reflectance spectra from the Munsell color atlas[9] of 16 complete hue sets of equal Munsell value and chroma. The spectra are measured in percent reflectance (scaled 0 to 1) at each nm from 400 nm to 700 nm and are represented by the matrix, $\mathbf{A}_{640\times301}$ in Eq. 3. We summarize the model in standard matrix notation[10] in the four equations below.

$$\mathbf{R}_{301\times3} \equiv \mathbf{U}_{301\times3}\mathbf{D}_{3\times3}\mathbf{V}^{\mathrm{T}}_{3\times3} \qquad [1]$$

$$\mathbf{P}_{301\times301} = \mathbf{U}_{301\times3}\mathbf{U}^{\mathrm{T}}_{3\times301} \quad [2]$$

$$\mathbf{S}_{640\times301} = \sqrt[3]{\mathbf{A}_{640\times301}}\mathbf{P}_{301\times301} \quad [3]$$

$$\mathbf{S}_{640\times301} \equiv \mathbf{M}_{640\times3}\Lambda_{3\times3}\mathbf{W}^{\mathrm{T}}_{3\times301} \quad [4]$$

Basically the model calculates an orthonormal projection matrix ($\mathbf{P}_{301\times301}$) for each set of observer cone sensitivity curves shown in Fig. 1 using the singular value decomposition in Eq. 1 followed by the matrix multiplication of Eq. 2. We then convert each spectrum into a linear transform of the cone sensitivity curves by multiplying the cube-rooted reflectance spectra by the projection matrix using Eq. 3. The projected reflectance spectra in matrix **S** represent what the receptors "see", in the literal sense that the reflectance spectra are now weighted combinations of the receptors. Another way of stating this is to say that the reflectance spectra are represented in the space of the photoreceptors.

We assume that the similarity of color appearance among the spectra is directly related to the similarity of the loci in the space of photoreceptors. The singular value decomposition of Eq. 4 provides a mathematical procedure for obtaining a three-dimensional Euclidean representation of the results. The output of the model comes in the form of two sets of three-dimensional coordinates. First, a set of coordinates that represent the locations of the 640 color chips in the three-dimensional matrix **M** of Eq. 4. Second, a set of coordinates that represent spectral color wavelength sensitivities in the three-dimensional matrix **W** of Eq. 4. For the color appearance of the chip locations, the first dimension corresponds to Munsell value going from black to white while the second and third dimensions correspond to chromaticity dimensions where angle around the origin represents hue and distance from origin represents chroma. For spectral wavelength sensitivities, the first dimension corresponds to achromatic sensitivity while the second and third dimensions correspond to spectral color sensitivity.

The results for the second and third dimensions are plotted in Fig. 2 which shows the estimated Munsell chip color appearance locations of the three observers together with the spectral color boundary for each type. The results for the normal observer were previously presented in Romney[3] and shown to correspond very closely to the Munsell color system. Calculations were carried out with the use of Mathematica[11]. The orientations of the three analyses were not identical in the output of Mathematica. We have rigidly rotated the protanomolous results to the same orientation as the normal observer results to facilitate comparisons.

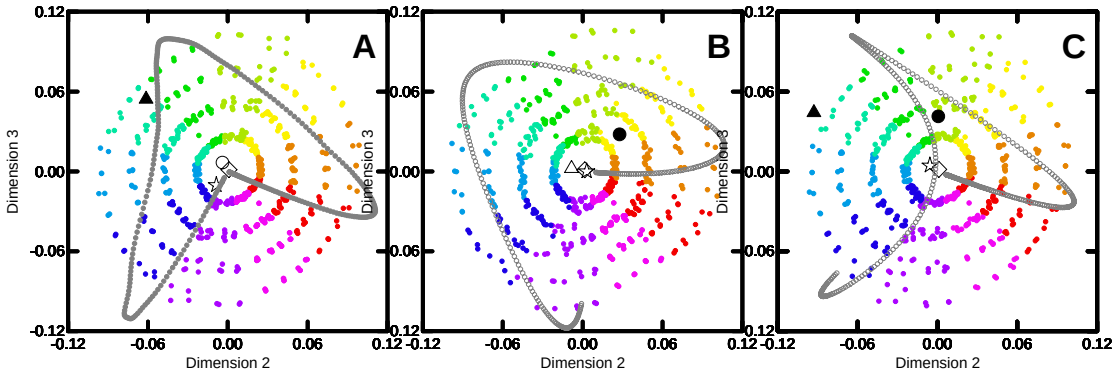

Fig. 2. Estimated color appearance of 640 Munsell chips and the spectral color boundary for each of the three observer types shown in Fig. 1. (A) Normal observer. (B) P3nm observer. (C) P0nm observer. The large symbols represent location of the curves presented in Fig. 3, when treated as color spectra, as follows: diamond=constant line, circle=normal observer, triangle=P3nm, and star=P0nm.

It may be seen that the two protanomalous types have somewhat distorted color vision in the placement of the location of the chips. Visually the placement of the chips for the two protanomalous types appear fairly similar compared to the normal observer. These predictions are derived mathematically and assume maximum efficient use of the observer's photoreceptors by the model. In practice the neurological machinery would not be expected to be so accurate. Even so, a visual examination of Fig. 2 would lead one to expect that all three observers may well pass a standard color test such as the Farnsworth-Munsell[12]. A multivariate measure of the similarity between the three dimensional coordinates of the estimated Munsell chip locations for the three observers is provided by the Stewart-Love redundancy index[13] and is above 0.99 for all comparisons among pairs of observers. A comparison of each of the observer results with the ideal Munsell conceptual coordinates also show an index above 0.99. The Stewart-Love index is a multivariate measure analogous to a squared correlation coefficient and measures the amount of variance accounted for among the variables. Thus the results shown in Fig. 2 reveal that the model provides estimates of the three protanomalous observers that are similar to each other as well as to the ideal Munsell conceptual locations.

In stark contrast to the similarity in the appearance of the Munsell color chips for the different observers, the shapes of the spectral color boundary are quite different from each other. There is a general similarity in the shape of the spectral color boundary for P3nm and the normal observer. On the other hand, P0nm has a color boundary radically different in shape from the other observers. The implications of these differences lead to critical differences in predictions for the three observer types. First, the model predicts different appearances of the metamers of the white point for each of the observers. Second, the model leads to some important differences in the predicted appearance of monochromatic spectral colors and complementary colors for the three observer types.

These results imply strong quantitatively precise predictions that we discuss in some detail in the following two sections. The predictions have not been tested but they

could be easily confirmed or disconfirmed using the techniques reported by Cognale *et al.*[14] as used to test predictions of the model in Romney[3]. The predictions are made on the assumption that the sum of the spectral color boundary curves are proportional to those reported by MacAdam[15] in his study of complementary colors. Our curves for a normal observer replicate MacAdam's finding of the number of watts of a spectrum complementary necessary for a neutral mixture with one watt of its complementary spectral color. We would have liked to estimate the Rayleigh matches reported by Neitz *et al*. but are unable to translate the units used in an anomaloscope into those used by MacAdam.

Predictions of the appearance of white metamers for different observer types

Fig. 3 shows a plot of the first spectral color boundary coordinate (exaggerated by multiplying by five) representing achromatic sensitivity for the three observer types together with a constant curve at 0.5. In the model the white point represents the illuminant and is assumed to be a flat line at some constant value for all observer types including those studied in this paper. In the model the first sensitivity coordinate represents sensitivity to achromatic light and is appropriately interpreted as a metamer of that observers white point[16]. We can check the model predictions for the curves of the three observers by treating them as reflectance spectra representing a color. We implement this by appending the curves to the 640 empirical spectra and applying the model to the whole set of 644 spectra. This will give us an estimate of their appearance for each of our three different observers. The resulting color appearance locations are plotted as large symbols in Fig. 2 and as spectrum curves in Fig. 3. The curve for the normal observer is the same curve that appears in Fig 3A in Romney[3].

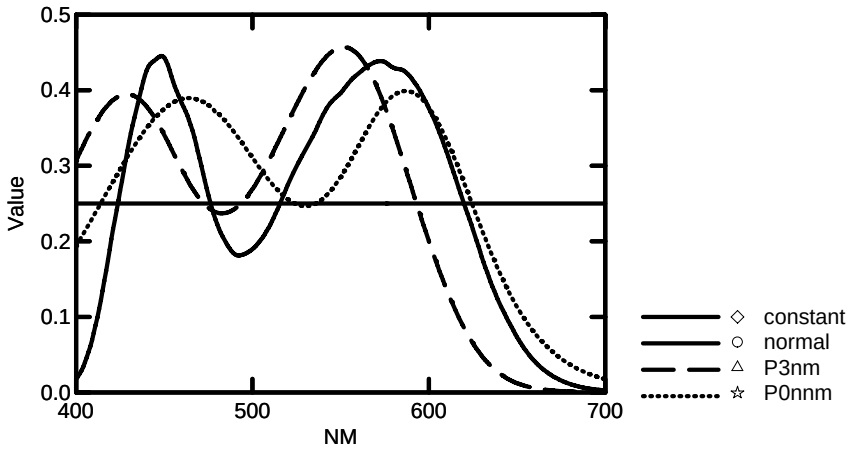

Fig. 3. Achromatic sensitivity curves multiplied by five (curved solid line is normal observer, dashed line P3nm observer, dotted line is P0nm observer) and a constant value curve of 0.5 (straight solid line).

For each observer the curved line in Fig. 3 represents the first achromatic sensitivity boundary curve for that observer and should appear achromatic. Therefore it should be a metamer of white for that observer but could appear as having a strong chromatic component for other observers. In Fig. 2 the color appearance of the curve for each observer should occur very near the center of their results, while the curve for the

other observers may appear with large chromatic components for that observer. An examination of the plots in Fig. 2 shows that in every plot the observers own achromatic sensitivity curve is very near the origin and is a close metemer of the constant white point. It is interesting to note that for each observer at least one of the other observer's achromatic points has a strong chromatic component, for example, the achromatic point for P3nm appears as a Munsell green with chroma=7 for the normal observer (Fig. 2A), the achromatic point for the normal observer appears as a Munsell yellow with chroma=3 for the P3nm observer (Fig. 2B), the achromatic point for the normal observer appears as a Munsell yellow-green with chroma=3 for the P0nm observer (Fig. 2C), and the achromatic point for P3nm observer appears as a Munsell blue-green with chroma =9 for the P0nm observer (Fig. 2C). Future empirical tests of these predictions will help confirm or disconfirm the model.

Model predictions for spectral colors and complementary colors

Perhaps the most unexpected and novel predictions made possible by the model are implicit in the spectral boundary differences among observers shown in Fig. 2. The distance from the origin of these curves at each wavelength defines the chromatic sensitivities of the respective observers and shown in Fig. 4A–C. In Fig. 4D–F the spectral color boundary is plotted for each observer type with the wavelength locations indicated by dotted vertical lines in Fig. 4A–C highlighted in red. The curve in Fig. 4A for the normal observer was first derived by MacAdam[15] over 80 years ago. His research involved complementary spectral colors and posed the question of how much of each was necessary to obtain a neutral achromatic white. He measured his curves in moments of watts of indicated wavelength required to neutralize one watt of a complementary wavelength. In MacAdam's derivation the short-wavelength peak is complementary to the second wavelength trough and the long-wavelength peak is complementary to the first wavelength trough, as is indicated by the two arrows in our Fig. 4A–B. Note that for the P0nm observer in Fig. 4C that the curve at first wavelength trough dips to 0 and hence no second arrow is drawn since there is no complementary color for the long-wavelength peak. The model predicts that a spectral light at 475 nm for the P0nm observer appears achromatic with apparent brightness proportional to the height of the dotted line in Fig. 3 at 475 nm. Note that for the normal observer there are complementary spectral colors at all wavelengths from 447 nm to 482 nm and for the P3nm observer there are complementary spectral colors for all wavelengths from 421 nm to 463 nm. In contrast, for the P0nm observer, there is only one area of complementary spectral colors around 421 nm.

The fact that the model predicts that a P0nm observer will lack a complementary spectral color for the long-wavelength peak sensitivity spectral color was completely unanticipated and provides a challenging empirical test for its validity. The finding of the shape of such an odd spectral color boundary for the P0nm observer places special emphases on the general location and prominence of the additive colors of red, green, and blue sensitivity lobes common to all observer types investigated to this time. For a normal observer there are no purple spectral colors, purple appearance arises from a mixture of red and blue spectral colors. For P0nm observers not only are there no purple spectral colors but there are no cyan spectral colors, cyan appearance arises from a mixture of green and blue spectral colors. A close examination of the placement of the

Munsell color chips and the spectral boundary colors in Fig 2 reveals that there are some large differences in the model predictions of the spectral colors among the three observer types. Using the plots as nomographs, one may easily calculate the Munsell hue predicted for each spectral color by the model. Future research will confirm or disconfirm the model predictions.

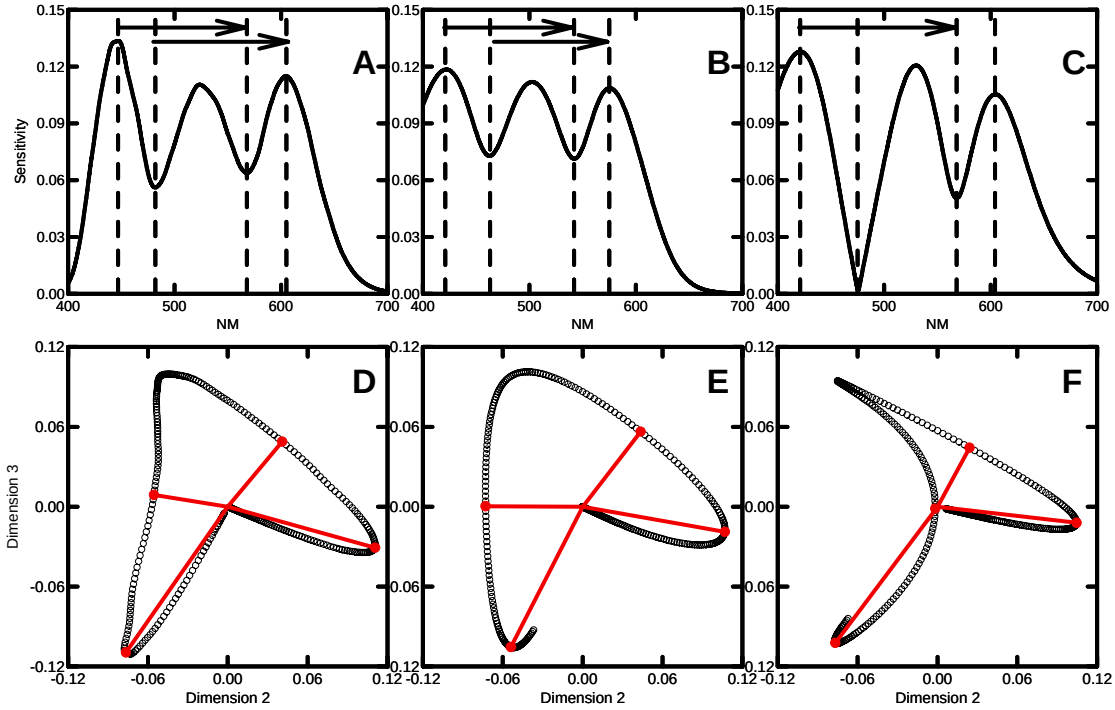

Fig. 4. The two chromatic components represented as chromatic sensitivities by wavelength and as spectral color boundaries. (A) Normal observer with peaks at 447 nm and 605 nm and troughs at 482 nm and 568 nm with arrows showing complementary sector. (B) P3nm observer with peaks at 421 nm and 575 nm and troughs at 463 nm and 542 nm with arrows showing complementary sector. (C) P0nm observer with peaks at 421 nm and 604 nm and troughs at 475 nm and 568 nm with single arrow indicating single narrow complementary color area. (D) Spectral color boundary for normal observer showing complementary sector bounded in red. (E) Spectral color boundary for P3nm observer showing complementary sector bounded in red. (F) Spectral color boundary for P0nm observer showing limited complementary pair area.

Discussion

There are three major observations about our results that need emphasis. First, the variant genetic types of protanomalous trichromats studied by Neitz *et al*. illustrate the need for a model that makes precise quantitative predictions about color appearance for any observer type. This paper is designed to fill that need by demonstrating how to estimate the color appearance for the Neitz *et al*. observers. To our knowledge no other color appearance model generalizes to all observer types. The ability to model any combination of genetically determined photoreceptor types is important because it is apparent that the laws of color combination and additivity such as Grassman's [17] laws are

specific to a specified observer type defined by known photoreceptor characteristics. Newton [18, 19] was entirely correct when he stated that monochromatic light beams have no inherent color. This research reinforces the fact that narrow and broad band light combinations can and do have different laws of combination and appearance for different observers, each different observer type follows a uniquely different Grassman's law. Even when the two eyes of a single individual express different photoreceptor genes, Grassman's law is different for each eye. For example, Alpern *et al*[20]. studied a unilateral tritanope who had normal color vision in one eye while the other lacked a short wavelength-sensitive cone and discovered that Grasmann's additivity law is grossly wrong for dichoptic matches.

Second, the rather remarkable similarity of the three observer types in terms of the appearance of the 640 Munsell chips requires some explanation, especially given the large discrepancies in the shape of the spectral color boundaries. Ordinary surfaces have reflectance spectra that are broad band and rather smooth with a single peak or trough and reflect at least some light at all wavelengths. The sensation of color is produced by the neural system comparing the responses of different classes of receptors. We assume that there are three receptor types that are broad-band and that at least two of them overlap at any wavelength. Under these constraints the orthonormal projection matrix produced by the application of Eqs. 1 and 2 will result in three independent and widely spaced peaks as observed in Figs. 4A–C. As outlined in previous papers[2,3] these peaks correspond to the additive colors of red, green, and blue. When combined, narrow-band spectral lights of these colors produce maximum color gamuts. Similarly visual systems with peak sensitivities in these areas, as for the observers shown in Fig. 2, would have trichromatic color vision. Even though the sensitivity peaks vary over a few tens of nm for different trichromatic observers, all have a modeled short wavelength peak somewhere around 420–460 nm, a medium wavelength peak around 510–550 nm, and a long wavelength peak around 590–630 nm nm as illustrated for the three observer types shown in Fig. 4. High definition color TV uses lights that peak near those of normal human observers at 447 nm, 525 nm, and 605 nm that produce the appearance of colors that encompass the gamut of the Musell chips.

Third, our results are consistent with numerous recent studies that demonstrate the enormous neural plasticity in the precise neural wiring from one individual to another. For example, Jacobs *et al*.[21] demonstrated the behavioral emergence of novel color vision in mice engineered to express human cone pigment. They interpreted this to indicate an inherent plasticity in the mammalian visual system that permits the emergence of a new dimension of sensory experience based solely on gene-driven changes in receptor organization. Wagner and Kröger[22] include a discussion of the adaptive plasticity during the development of color vision in fish, humans, and other organisms. Even more remarkable are changes from dichromatic to trichromatic color vision in adult primates when provided with an additional receptor through gene therapy as reported by Mancuso *et al.*[23]. It is clear from an examination of Fig. 2 that no rigid and specifically detailed genetic program could apply in the same way for the genetic combinations of photoreceptor types to produce such similar appearance percepts of the Munsell chips. The modeling of distinctive protanomalous trichromat observer types adds additional confirmation of the necessity for a neural plasticity that facilitates the computation of

optimal color vision utilizing whatever set of photoreceptor sensitivities an organism inherits (or obtains through gene therapy).

References

1. Neitz, J., Neitz, M., He, J.C. & Shevell, S.K. Trichromatic color vision with only two spectrally distinct photopigments. *Nat. Neurosci.* **2**, 884-888 (1999).
2. Romney, A.K. & Chiao, C.-C. Functional computational model for optimal color coding. *P. Natl. Acad. Sci. USA* **106**, 10376-10381 (2009).
3. Romney , A.K. Modeling color vision with and without and observer. *J. Opt. Soc. Am. A* **under review**, 1-9 (2009).
4. Stockman, A. & Sharpe, L.T. The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Res.* **40**, 1711-1737 (2000).
5. Govardovskii, V.I., Fyhrquist, N., Reuter, T., Kuzmin, D.G. & Donner, K. In search of the visual pigment template. *Visual Neurosci.* **17**, 509-528 (2000).
6. Shevell, S.K. & Chang He, J.I. The Visual Photopigments of Simple Deuteranomalous Trichromats Inferred from Color Matching. *Vision Res.* **37**, 1115-1127 (1997).
7. He, J.C. & Shevell, S.K. Variation in color matching and discrimination among deuteranomalous trichromats: Theoretical implications of small differences in photopigments. *Vision Res.* **35**, 2579-2588 (1995).
8. Shevell, S.K., Sun, Y. & Neitz, M. Protanomaly without darkened red is deuteranopia with rods. *Vision Res.* **48**, 2599-2603 (2008).
9. Munsell Color Company, I. *Munsell Book of Color. Matte Finish Collection* (Munsell, 1976).
10. Strang, G. *Linear Algebra and Its Applications* (Harcourt, Brace, Jovanovich, San Diego, CA, 1988).
11. Wolfram, S. *The Mathematica Book* (Cambridge University Press, 1999).
12. Farnsworth, D. The Farnsworth-Munsell 100-Hue and Dichotomous Tests for Color Vision. *J. Opt. Soc. Am.* **33**, 568-574 (1943).
13. Stewart, D. & Love, W. A general canonical correlation index. *Psychol. Bull.* **70**, 160-& (1968).
14. Cognale, M.A., Webster, A.R. & Fong, A.Y. Application of digital micromirror defices to vision science: shaping the spectrum of stimuli. *Proceedings of SPIE* **7210**, 1-7 (2009).
15. Macadam, D.L. Photometric relationships between complementary colors. *J. Opt. Soc. Am.* **28**, 103-103 (1938).
16. Macadam, D.L. *Color Measurement: Theme and Variations* (Springer-Verlag Berlin Heidelberg New York, 1981).
17. Grassmann, H.G. Theory of Compound Colors (1853). In *Sources of Color Science*, MacAdam, D. L., ed. (MIT Press, Cambridge, MA, 1970).

18. Newton, I. New Theory about Light and Colors (1672). In *Sources of Color Science*, MacAdam, D. L., ed (MIT Press, Cambridge, MA 1970).
19. Newton, I. Opticks (1704). In *Sources of Color Science*, MacAdam, D. L., ed. (MIT Press, Cambridge, MA, 1970).
20. Alpern, M., Kitahara, K. & Krantz, D.H. Perception of color in unilateral tritanopia. *J. Physiol.-London* **335**, 683-697 (1983).
21. Jacobs, G.H., Williams, G.A., Cahill, H. & Nathans, J. Emergence of novel color vision in mice engineered to express a human cone photopigment. *Science* **315**, 1723-1725 (2007).
22. Wagner, H.J. & Kroger, R.H.H. Adaptive plasticity during the development of colour vision. *Prog. Retin. Eye Res.* **24**, 521-536 (2005).
23. Mancuso, K., *et al.* Gene therapy for red-green colour blindness in adult primates. *Nature* **advance online publication** (2009).