# The Challenge of our Known Unknowns

*Robert W. G. Hunt, Color Consultant, Salisbury, England; The University of Leeds, England*

## Abstract

*Although much is known about color vision and imaging, there are still important gaps in our knowledge, and the possible impact of these gaps needs addressing. Such topics include the following. Why are color-matching functions based on matches on white, and on matches on spectral colors, different, and what are the implications of this on color technology and imaging? Why does sharpening color-matching functions lead to better chromatic adaptation transforms? Why do the unique hues occur where they do in color space? How are the rods inhibited at high levels of illumination? Why do bluer whites look whiter than neutral whites of the same reflectance, and why is this also true of blacks? How can predicting the color rendering properties of white LEDs be improved? How can the use of true luminance signals be achieved? How can displays using a luminance signal be engineered?*

## Introduction

It was Donald Rumsfeld who famously said[1]:

As we know, there are known knowns.
There are things we know we know.
We also know there are known unknowns.
That is to say, we know there are some
things we do not know.
But there are also unknown unknowns,
the ones we don't know we don't know.

So what are some of the important known unknowns in color science and imaging? In color science, known unknowns include the following. Reasons for the difference between the maximum-saturation and the Maxwell-method for determining color-matching functions; why sharpened color-matching functions in chromatic adaptation transforms give improved results; the physiological basis for the positions of the unique hues in color spaces; the way in which rod responses are inhibited by cone activity; the reason why bluer whites look whiter; and an effective way of evaluating the color-rendering properties of light-emitting-diode light sources. In color imaging, known unknowns include the following. The most effective method of using a true luminance signal; and a practical way of providing displays that use chrominance and luminance signals.

## Color Science

### *Maxwell color-matching functions*

Two different methods of obtaining color-matching functions have been described in the literature. The Maximum-Saturation method, in which each wavelength of the spectrum is matched with a mixture of red, green, and blue primaries (one, or occasionally two, of which may have to be added to the spectral color being matched); and the Maxwell[2,3] method in which white light is matched by mixtures of a wavelength of the spectrum plus two of the red, green, and blue primaries. When comparison of the results obtained by the two methods is made, significant differences are found[4,5]. When shown on chromaticity diagrams, the results obtained by all investigators show the Maxwell spectrum locus of greenish colors to be placed further out from the white region than with the Maximum-Saturation method. The magnitude of the effect found is different in the investigations reported by different workers, but it is always of the same character. An example derived from the observations made by Wyszecki[6] is shown in Figure 1.

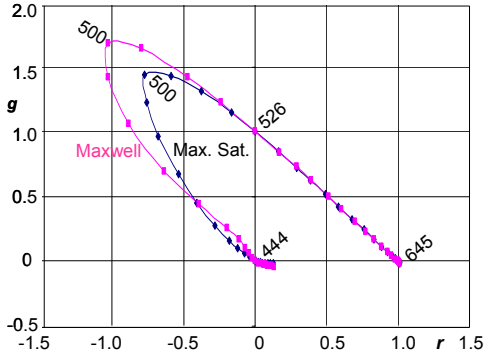


***Figure 1.*** *Chromaticities of matches on spectral colors, obtained by the Maximum-Saturation method, and by the Maxwell method. Results obtained by Wyszecki, 1978.*

What is the cause of the different results obtained by the Maximum-Saturation and by the Maxwell methods? Wyszecki found that a similar displacement in the chromaticity diagram occurred when the Maxwell method was used at 100,000 and 1000 trolands of retinal illumination; this suggests that changes in adaptation can cause such displacements, and the adaptation is very different in the Maximum-Saturation and the Maxwell methods. The rate of bleaching of the retinal photo-pigments will therefore also be different. But, at normal levels of retinal illumination, metameric matches are not upset by changes in adaptation[7], and this implies that the shapes of the cone spectral sensitivity curves are constant. However, it is known that, when rhodopsin is bleached, it can produce metarhodopsin II which has a yellow colour; it is therefore possible that, when cone pigments are bleached, they also produce yellowish decomposition photo-products, and these could change the effective spectral sensitivities of the receptors into one constant set for the Maximum-Saturation method and a slightly different set for the Maxwell method. Support for this view is perhaps provided by the fact that it has been reported that 'observed blue tristimulus values were often much larger than those calculated'[5]; in this case, the calculations were based on Maximum-Saturation color-matching functions, and the matches were on less saturated stimuli. If the less saturated stimuli produced more decomposition products, the blue tristimulus values would have to increase. Furthermore, the variation in the magnitude of the blue shifts found in different investigations could be caused by the observers having had different adaptation-exposure prior to the experiments being carried out. The several minutes adaptation, typically used prior to starting observations, might not be enough to erase the effects of the previous hours. There could thus be a slow (several hours)

adaptation effect caused by decomposition products, in addition to the fast (a few minutes) adaptation caused by electronic amplification of the signals from the retina.

It is, perhaps, worth recalling the 'positive blue phenomenon' reported by Wright[8]. He found that an unusually large amount of blue was required in the matches made during the recovery after adaptation to colored stimuli. The reason for this was not clear, but it does provide another example of an increase in blueness associated with changes in adaptation.

If the differences between the Maximum- Saturation and the Maxwell color-matching data are caused by adaptation, which should be used? The colors for which colorimetry is used in practice are seldom of near-spectral purity, and are much more often of quite low purity. So perhaps practical color-matching functions should be based on Maxwell data. It would be interesting to know whether such a practice would reduce the departures from additivity that have been reported.

### *Why does sharpening color-matching functions lead to better chromatic adaptation transforms?*

Following the earlier example by Lam[9], several more recent investigations have led to chromatic adaptation transforms being based, not on likely cone responses, but on matrixed responses giving sharpened color-matching functions[10,11,12]. The use of these transforms usually results in better predictions of chromatic adaptation, and this type of transformation has been used in the color appearance model CIECAM02. It is interesting to ask why such transformations result in better predictions. One implication is that correction for illumination takes place, not in cone space, but rather in a narrowed cone space. But where would this occur? Presumably somewhere between the cones themselves and the cells where color-difference signals are formed.

### *Why do the unique hues occur where they do in color space?*

Of all the characteristics of color perception, the existence the four unique hues, red, green, yellow, and blue, is one of the most striking. That their existence depends on signals derived from differences between cone responses has long been postulated from psychophysical studies, and the existence of such signals in various species has more recently lent support to this view. But the reasons for the exact location in color space of these unique hues are not yet clear.

The curvatures of the unique-hue loci on chromaticity diagrams indicate that the criteria for the unique hues occur after the linear stage of color vision, and hence after the absorption of the light in the cones. The discontinuities of the red-green, and of the yellow-blue, loci at the achromatic point, indicate that the criteria for unique red and unique green are different from one another, and that those for unique yellow and unique blue are also different from one another.

In the CIECAM02 color appearance model, the predictor for hue is based on the unique-hue criteria proposed in earlier models of color appearance[13,14,15]. These criteria are based on the following differences between the cone responses, $\rho$ (for the Long-wavelength cones), $\gamma$ (for the Medium-wavelength cones), and $\beta$ (for the Short-wavelength cones):

| | |
|---|---|
| Unique red | $C_1 = C_2$ |
| Unique green | $C_1 = C_3$ |
| Unique yellow | $C_1 = C_2/11$ |
| Unique blue | $C_1 = C_2/4$ |

where $C_1 = \rho - \gamma$
$C_2 = \gamma - \beta$
$C_3 = \beta - \rho$

These criteria give very good predictions of the unique hue loci, as shown in Figure 2, but they are based on the Hunt-Pointer-Estevez (HPE) cone spectral-sensitivity curves, of which the the ρ curve is more separated from the γ curve, peaking at about 580 nm instead of at about 560 nm as is the case for the more widely accepted Smith and Pokorny[16] or Stockman and Sharpe[17] curves. The matrix used to derive the HPE cone curves is

$$\rho = 0.38971X + 0.28898Y - 0.07868Z$$
$$\gamma = -0.22981X + 1.18340Y + 0.04641Z$$
$$\beta = 0.00000X + 0.00000Y + 1.00000Z$$

The following matrix gives curves that approximate the Smith and Pokorny or Stockman and Sharpe curves more closely:

$$\rho = 0.23X + 0.80Y - 0.03Z$$
$$\gamma = -0.55X + 1.45Y + 0.10Z$$
$$\beta = 0.00X + 0.00Y + 1.00Z$$

However, as shown in Figure 3, the simple criterion for unique red no longer gives good prediction of the unique red locus. So is this another indication that some sharpening of the cone responses takes place, in this context before the unique hues are established, or do the hue criteria in the model not have a physiological basis? These are important unknowns, answers to which call out to be established. The factors 11 and 4 in the predictors for yellow and blue are also as yet without a physiological basis.

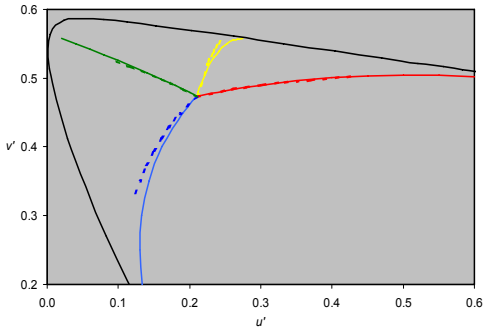


***Figure 2.*** *Unique hue loci: full lines, predictions using HPE cone curves; broken lines, NCS experimental results.*

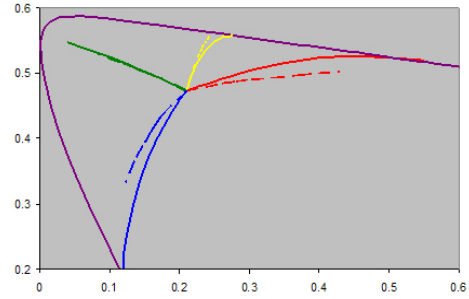


***Figure 3.*** *Same as Figure 2, but using, for the predictions, cone curves approximating those of the Smith and Pokorny cone curves.*

### How are the rods inhibited at high levels of illumination?

At mesopic levels of illumination, both the cones and the rods are active, and at photopic levels some significant rod activity can occur in large fields of view, as found by Stiles in his work on which the CIE 1964 Standard Colorimetric Observer was established[18]. However, the possibility of rod activity is ignored in practical colorimetry, when using both the $2^0$ and the $10^0$ Standard Observers. In imaging, the rods are also ignored, although in this case the field sizes of importance are usually quite small. The fact that colorimetry works well, both in general applications and in imaging, indicates that, at usual photopic levels of illumination, the rods are inactive. This is generally explained by saying that, in these conditions, the cones inhibit the rods. Indeed, if the rods were not inhibited they would add a desaturating achromatic response that became larger and larger as the level of illumination increased; this clearly does not happen. But just how this inhibition takes place is an important unknown.

### Why do bluer whites look whiter than neutral whites of the same reflectance, and why is this also true of blacks?

When two whites having the same reflectance factor are compared, if one is bluer than the other it looks whiter. Does this mean that there is something wrong with our $V(\lambda)$ functions? It would seem not to be so; because in the CIE Whiteness Index, and in other whiteness formulae, rather than using a different $V(\lambda)$ function, it is found necessary to add a factor that represents increasing whiteness as the chromaticity becomes bluer. Moreover, it has also been known for many years by launderers that white materials can be made to look whiter by adding a small amount of blue dye; in this case the reflectance factor is actually decreased, but the effect is to increase the whiteness. It has also been reported recently[19] that blacks of similar reflectances appeared blacker if their chromaticities are in the blue direction. It has sometimes been suggested that the reason for bluish whites appearing whiter is because, when white materials deteriorate, they usually become yellower; even if this is true for whites, it seems less plausible for blacks. Here is another interesting unknown.

### How can predicting the color rendering properties of white LEDs be improved?

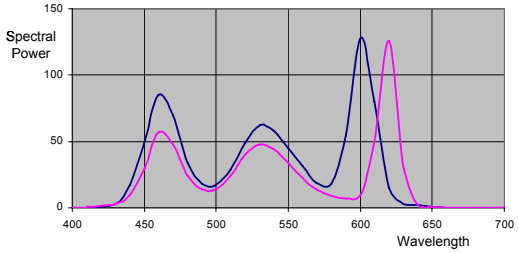


***Figure 4.*** *Spectral power distributions for two white LED sources: Source A, black line; Source B, magenta line.*

The CIE Color Rendering Index is based on comparing the rendering of eight Munsell samples under the illuminant considered with the rendering of the same samples under either a CIE D illuminant or a Planckian radiator. It has been found that this index performs poorly when it is used for evaluating the rendering of white LED sources. An example is given in Figure 4. These two light sources have the same correlated color temperatures (about 5000K). Source A has a CIE Color Rendereing index, $R_a$, equal to 82, while for Source B it is 71; the value for the reference illuminant is 100. This implies that Source A has better color rendering than Source B. But in fact the reverse is true; Source B renders saturated reddish colors much closer to the reference illuminant, D50, than Source A.

| Spectral band (nm) | Band luminance for D50 | Band luminance for A | Ratio of band luminances | 1 - ratio | Band deviation % | Excess over tolerance |
|---|---|---|---|---|---|---|
| 400-455 | 0.585 | 0.537 | 0.917 | 0.083 | -8 | 0 |
| 400-510 | | | | | -18 | 13 |
| 455-510 | 9.52 | 6.80 | 0.715 | 0.286 | -29 | 19 |
| 455-540 | | | | | -3 | 0 |
| 510-540 | 21.9 | 26.7 | 1.221 | 0.221 | 22 | 12 |
| 510-590 | | | | | -1 | 0 |
| 540-590 | 44.3 | 33.3 | 0.752 | 0.248 | -25 | 15 |
| 540-620 | | | | | 39 | 34 |
| 590-620 | 15.8 | 32.1 | 2.032 | 1.032 | 103 | 93 |
| 590-760 | | | | | 5 | 0 |
| 620-760 | 7.94 | 0.534 | 0.067 | -0.933 | -93 | 83 |

Sum of excesses 269

Figure of merit = 1024 – 269 = 755

***Table 1.*** *Use of the spectral band method of assessing color rendering with Source A. The tolerances are ± 10% for single bands and ± 5% for the average percentage difference in all pairs of contiguous bands.*

| Spectral band (nm) | Band luminance for D50 | Band luminance for B | Ratio of band luminances | 1 - ratio | Band deviation % | Excess over tolerance |
|---|---|---|---|---|---|---|
| 400-455 | 0.585 | 0.519 | 0.887 | 0.113 | -11 | 1 |
| 400-510 | | | | | -16 | 11 |
| 455-510 | 9.52 | 7.63 | 0.801 | 0.199 | -20 | 10 |
| 455-540 | | | | | 12 | 7 |
| 510-540 | 21.9 | 31.6 | 1.442 | 0.442 | 44 | 34 |
| 510-590 | | | | | 7 | 2 |
| 540-590 | 44.3 | 31.2 | 0.704 | -0.296 | -30 | 20 |
| 540-620 | | | | | -4 | 0 |
| 590-620 | 15.8 | 19.2 | 1.218 | 0.218 | 22 | 12 |
| 590-760 | | | | | 23 | 18 |
| 620-760 | 7.94 | 9.86 | 1.241 | 0.241 | 24 | 14 |

Sum of excesses 129

Figure of merit = 1024 – 129 = 895

***Table 2.*** *Same as Table 1, but for Source B.*

At the time when the CIE Color Rendering Index was devised, there was an alternative method[20,21,22] based on spectral bands as shown in Tables 1 and 2. Using this method Source A has a value of 755, and Source B a value of 895, compared to a value of 1024 for the reference illuminant. This is because the

spectral power of Source B peaks more nearly at Thornton's prime wavelengths, 450, 530, and 610 nm than is the case for Source A. It is not known whether this spectral band method gives satisfactory results for all sources, but there is a rather urgent need to provide a measure of the color rendering of sources that covers LEDs and all other sources used for illumination.

## Color Imaging

### *How can the use of true luminance signals be achieved?*

Because of gamma correction, the luminance signal used in broadcast television is not a true luminance signal, and, for saturated colors, a significant amount of luminance is actually carried by the chrominance signals[23]. This is why, in high-definition broadcast television, the chrominance signals are reduced in bandwidth to only one-half, instead of to one-quarter, of the luminance signal. True luminance signals can be obtained by using cameras that have a luminance sensor, together either with just a red and a blue sensor, or with a red, a green, and a blue sensor (the latter arrangement requiring less signal processing)[24]. The best way to implement these ideas is still open to debate At present quite a large penalty is incurred by not using a true luminance signal.

### *Displays using luminance signals*

Conventional electronic displays use the same number of pixels for red, green, and blue; that is 3*N* pixels, if *N* is the number of composite pixels displayed. A display using separate luminance and chrominance would need only *N* pixels for the luminance and *N*/16 pixels for each of the three colors, making a total of $N + (3/16)N$ or $(19/16)N$. This is assuming that a true luminance signal is used, thus allowing a reduction in chrominance to one quarter of the bandwidth in the horizontal direction and a similar reduction of the definition in the vertical direction. A problem with this approach is that the chrominance display needs constant luminance at all chromaticities; this restricts the luminance to those of the blues which are very low. However, the display could consist of a low definition part that has correct chrominance but maximum luminance in each group of 4 x 4 pixels, and a high-definition luminance correcting panel consisting of transmitting elements that reduce the luminance of each pixel to its correct value. In this way the total number of pixels is still $(19/16)N$ but the restriction of the luminance is avoided[25,26,27].

## Acknowledgements

I am grateful to Dr Michael Pointer for helpful discussions in the preparation of this paper.

## References

[1] D. Rumsfeld. Feb. 12, 2002, Department of Defense news briefing.

[2] J.C.Maxwell. On the theory of compound colours, and the relations of the colours of the spectrum. *Phil. Trans. Roy. Soc. (London)* **150,** 57-84 (1860).

[3] J.C.Maxwell. Theory of compound colours, and the relations of the colours of the spectrum. *Proc. Roy. Soc.* (London) **10,** 404-409 (1860).

[4] B.H.Crawford. Colour matching, and adaptation. *Vis. Res.* **5,** 71-78 (1965).

[5] R.D.Lozano and D.A.Palmer. Large-field colour matching and adaptation. *J. Opt. Soc. Amer.* **58,** 1653-1656 (1968).

[6] G.Wyszecki and W.S.Stiles. *Color Science.2nd ed.,* pages 379 to 392, Wiley (1982).

[7] W.D.Wright.*Researches on Normal and Defective Colour Vision,* page 148, Henry Kimpton, London (1946).

[8] W.D.Wright.*Researches on Normal and Defective Colour Vision,* pages 252-254, Henry Kimpton, London (1946).

[9] K.M.Lam. Metamerism and colour constancy. Ph.D. Thesis, University of Bradford, 1985

[10] G.D.Finlayson and S.Susstrunk. Performance of a chromatic adaptation transform based on spectral sharpening. *CIC* **8**, 49-55 (2000).

[11] G.D.Finlayson and S.Susstrunk. Spherical sampling and color transformations. *CIC* **9**, 321-325 (2001).

[12] S.Susstrunk and G.D.Finlayson. Evaluating chromatic adaptation transform performance. *CIC* **13**, 75-78 (2005).

[13] R.W.G.Hunt, A model of colour vision for predicting colour appearance. *Col. Res. Appl.,* **7,** 95-112 (1982).

[14] R.W.G.Hunt. *Measuring Colour, 3rd ed.,* pages 208-237, Fountain Press (1998).

[15] R.W.G.Hunt. *The Reproduction of Colour, 6th ed.,* pages 596-616, Wiley (2004).

[16] V.C.Smith and J.C.Pokorny. Spectral sensitivity of colorblind observers and the cone pigments. *Vis. Res.* **12,** 2059-2071 (1972).

[17] A.Stockman. Physiologically based color-matching functions. *CIC* **16**, 1-5 (2008).

[18] W.S.Stiles and J.M.Burch. N.P.L. colour-matching investigation: final report (1958). *Optica Acta,* **6,** 1-26 (1959).

[19] R.Clonts and R.Shamey. Perceptual assessment of blackness. *ISCC Special Topics Meeting: Black and White,* pages 20-21 (2008).

[20] B.H.Crawford. *Brit. J. Appl. Phys.,* **14**, 319-328 (1963).

[21] B.H.Crawford. Colour-rendering tolerances and the colour-rendering properties of light sources. *Trans. Illum. Eng. Soc.,* **28**, 50-64 (1963).

[22] British Standard 950: 1967. Artificial daylight for the assessment of colour. Parts I and II (1967).

[23] R.W.G.Hunt. *The Reproduction of Colour ,6th ed.,* pages 401-402, Wiley (2004).

[24] R.W.G.Hunt. *The Reproduction of Colour, 6th ed.,* pages 368-372, Wiley (2004).

[25] G.Dispoto, S.Susstrunk, and R.W.G.Hunt. Display method and system using transmissive and emissive components. U.S.Patent 20060061538 A1 (2006).

[26] R.W.G.Hunt, S.Susstrunk, J.M.Holm, and G.J.Dispoto Display system and method using a projector and a reflective display. US Patent 7,466,309 (2008).

[27] Y.Kusakabe, M.Kanazawa, Y.Nojiri, M.Furuya, and M.Yoshimura. A YC-separation-type projector: High dynamic range with double modulation. *Journal of the ASID* **16/2**, (2008).