The Linearity of Colour Appearance Mechanisms

Sophie M Wuerger, Centre for Cognitive Neuroscience, School of Psychology, University of Liverpool, L697AZ, UK

Abstract

Our aim was to characterise the higher-order chromatic mechanisms that yield the four unique hues: red, green, yellow and blue. Our results are consistent with the hypothesis that all four unique hues are generated by mechanisms that are linear in XYZ space. Furthermore, the variability between observers is relatively small when expressed in terms of perceptual errors and, as a consequence, the same linear model (same weights) fits the data for all observers. Given the linearity of these higher-order colour appearance mechanisms and the consistency across observers, these appearance judgments may be suitable to derive a standard observer model for colour appearance. Our results add further weight to the idea that the colour vision system in adult humans is able to recalibrate itself based on prior visual experience.

Introduction

When human observers are asked to adjust a coloured light such that it appears neither red nor green, or such that is appears neither yellow nor blue, most colour-normal observers have no difficulty in making these adjustments [1] and these colour appearance judgments are not influenced by culture and language [2], or by age [3, 4]. Furthermore, the variability in these colour appearance judgments is not related to the variability in chromatic discrimination sensitivity [5]. All these findings suggest that there is something very fundamental about these four attributes: redness, greenness, yellowness and blueness. The aim of this study is twofold: (1) to establish whether these colour appearance mechanisms (yielding these four colour attributes) are linear mechanisms in XYZ space and (2) to evaluate the variability across observers. To evaluate the goodness of fit of a linear model holds we use a perceptual error metric (CIE 1994).

Methods

Task

We obtained unique hue settings from 18 colour-normal observers (confirmed with the Colour Vision Test from Cambridge Research Systems) for a wide range of luminance and saturation levels, using a hue-selection task. For instance, to obtain the settings for unique red, the observer had to judge which patch contained neither yellow nor blue. In each trial an annulus of 12 coloured disks covering different shades of red was presented (see Figure 1). The observer made a selection by clicking with the mouse on the coloured patch, which contained neither yellow nor blue. Unique green settings were obtained similarly: an annulus of greenish patches was presented and the observer had to select that patch that contained 'neither yellow nor blue'. Unique yellow (blue) settings were obtained by presenting an annulus of yellowish (bluish) patches and asking the observer to select that patch contains 'neither red nor green'. Each individual disk had a diameter of 1.5 degrees of visual angle; the radius of the annulus was 5 degrees and was centered at the midpoint of the screen; on each trial the colour was assigned randomly to a particular patch on the annulus. All observers found the task easy and did not need require any further instructions or any explanations what 'neither red not green' (or 'neither yellow nor blue') meant.

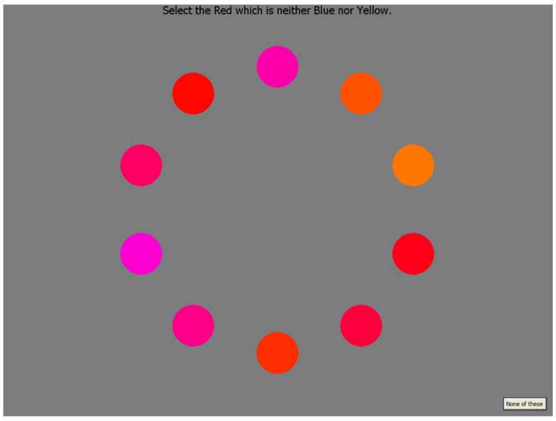

Figure 1*. On each trial an annulus of coloured patches was presented on a grey background. To obtain the settings for unique red, for instance, the patches were all reddish and the observer had to judge which patch contained 'neither yellow nor blue'. The observer selected the most appropriate colour patch by clicking with the mouse on the patch. There was no time limit and the observer was encouraged to move his/her eyes freely.*

Colour Space

For stimulus selection we used the HSV (Hue-Saturation-Value) colour space, since we attempted to keep the saturation and the luminance level approximately the same for all the colour patches presented on a particular trial; this facilitates the task of the observer to select the most appropriate hue. The HSV space is scaled such that it makes use of the entire monitor gamut. 'Value' (which is loosely related to luminance) can range from 0 to 1; 0 is black and 1 is white (for an achromatic colour). Saturation can range from 0 to 1 and refers to the amount of grey in a particular colour; a saturation of 0 indicates a grey colour; a saturation of 1 refers to a fully saturated colour. Hue is specified as an angle ranging from 0 to 360 deg. In each session all four unique hues were determined at different saturation and luminance levels. Based on preliminary experiments, we used saturation levels ranging from 0.2 to 1.0, in steps of 0.05, and value levels ranging from 0.3 to 1.0, again in steps of 0.05. These value levels resulted in luminances ranging from 2 cd/m^2 to 130 cd/m^2, with an average luminance of about 42 cd/m^2. At a particular trial, a specific combination of saturation and level was used; the order of presentation was randomised.

For the data analysis, the colour coordinates were converted from HSV to XYZ space [6]. The parameters for the linear model (cf Eq. 1-4). were estimated in XYZ space. To evaluate the goodness of fit and the variability between observers, perceptual errors were computed in L*a*b* space, which is an approximately uniform colour space.

Apparatus

All stimuli were presented on a CRT screen of a DELL monitor (DELL P790). Linearised look-up tables were produced by measuring the CRT light outputs with a spectroradiometer (SpectraScan PR650; PhotoResearch). The background was always grey with a mean luminance of 43 cd/m^2 and with chromaticity co-ordinates x=0.282 and y= 0.307. The observers were seated in a darkened room 1 m away from the monitor and adapted to the grey background for at least 5 minutes. The stimuli were presented continuously until the observer responded. There was no time limit for the response and the observers were encouraged to move their eyes freely. Each observer made at least 80 selections for each unique hue. Altogether, for each unique hue we obtained 1616 data points.

The Linear Model

The assumption underlying hue cancellation judgments [7] is that each unique hue is generated by silencing a chromatically opponent mechanism. That is, all colours that are judged as 'unique red' appear 'neither yellow nor blue', hence silencing a yellow-blue (YB) mechanism. If unique red (R) is generated by silencing an YB mechanism and if this mechanism is linear in XYZ space, then we can write:

$$\alpha_R * \Delta X + \beta_R * \Delta Y + \gamma_R * \Delta Z = YB_R = 0 \qquad (1)$$

All colour patches were presented on a grey background the observers were adapted to. Since colour appearance depends mainly on the incremental signal with respect to this background, the linear model is formulated in terms of incremental (or decremental) XYZ co-ordinates, denoted with ΔX, ΔY, ΔZ. Equation 1 defines a plane (through the origin) in XYZ space. The normal vector $(\alpha_R, \beta_R, \gamma_R)$ characterises the yellow-blue mechanism (YB_R) that is silenced by all the (unique red) colours on this plane. We can derive an analogous equation for unique green:

$$\alpha_G * \Delta X + \beta_G * \Delta Y + \gamma_G * \Delta Z = YB_G = 0 \qquad (2)$$

For unique yellow we assume that a particular red-green mechanism (RG_Y) is silenced since unique yellow is obtained by selecting a yellowish light that appears 'neither red nor green'. The null plane for this opponent RG mechanism is therefore defined as:

$$\alpha_Y * \Delta X + \beta_Y * \Delta Y + \gamma_Y * \Delta Z = RG_Y = 0 \qquad (3)$$

The vector $(\alpha_Y, \beta_Y, \gamma_Y)$ is orthogonal to this plane and characterises the red-green mechanism (RG_Y), which is silenced by all colours on this plane. An analogous null plane can be derived for unique blue; the corresponding normal vector characterises the red-green mechanism (RG_B), which is silenced by all colours on this null plane:

$$\alpha_B * \Delta X + \beta_B * \Delta Y + \gamma_B * \Delta Z = RG_B = 0 \qquad (4)$$

Now we ask the following questions: (1) Does a linear model defined by Eq. 1-4 fit the data? (2) Do the same weights predict the unique hue settings for all observers implying that observers have very similar colour appearance mechanism?.

Results and Discussion

Unique Hue Settings in the CIE xy diagram

The unique hue settings were obtained for a range of luminance and saturation levels. These settings obtained in HSV space were then converted to XYZ space and each unique hue selection is a point in this three-dimensional XYZ space. To visualise the loci of the unique hues, we first present our data in the two-dimensional CIE diagram (Figure 1); all the analysis however is performed in XYZ space. In figure 1, red symbols (circles) refer to the settings for 'unique red' and green symbols (squares) to 'unique green'; yellow symbols (upward pointing triangles) denote the data for 'unique yellow' and blue symbols (downward pointing triangles) indicate the loci of 'unique blue'. The dark outline indicates the monitor gamut.

Figure 2 demonstrates an important and well-known property of the unique hues: whereas yellow and blue are collinear in the xy diagram and lie on a line through the grey origin (adapting background colour), the settings for red and green do not lie on a line through the origin [8, 9].

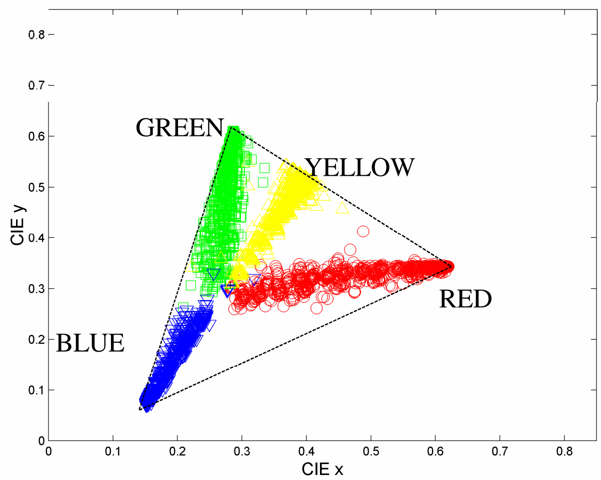

***Figure 2**. The unique hue settings are shown in the CIE xy diagram. The unqiue hue settings were obtained at a large range of luminance and levels (which cannot be seen in the xy diagram). The red circles denote the settings for red, the yellow upward-pointing triangles for yellow, the green squares for green, and the blue down-ward pointing triangles for blue. The dark outline indicates the monitor gamut.*

Linearity of the unique-hue mechanisms

The first aim of this study was to test whether the unique hue mechanisms are linear transformations of XYZ space. To evaluate linearity we determined the best-fitting linear model and evaluated the deviations of the data from it.

For each of the four data sets (RED, GREEN, YELLOW, BLUE) we estimate the coefficients α, β, γ (as defined in Eq. 1 – 4) by minimizing the Euclidean distances between the unique hue settings (in XYZ space) and the theoretical plane ('Orthogonal Distance Regression' [10]). The coefficients α, β, γ define the vector which is orthogonal to the plane (normal vector): the dot product between the normal vector and any point on this plane is zero. This normal vector therefore corresponds to the chromatic mechanism which is silenced by all the settings for a particular hue. For instance, all the reddish colours that are considered as 'neither yellow nor blue' (that is, all unique red settings) lie on a particular surface in XYZ space. The linear model assumes that this surface is a plane. The normal vector for the 'unique red plane' hence defines the chromatic mechanism (in XYZ space) that gives zero response to all unique red settings (Eq. 1). The length of this normal vector is arbitrary and the coefficients are normalised such that $\alpha^2 + \beta^2 + \gamma^2 = 1$. For each unique hue we fitted n=1616 data points. An equivalent method for finding the coefficients (α, β, γ) is a principal component analysis (PCA). The last eigenvector (explaining the least variance) is the normal vector.

Using orthogonal distance regression we obtained the following coefficients α, β, γ. For RED: 0.4662, -0.8483, 0.2510, for GREEN: -0.8868, 0.3716, 0.2747, for YELLOW: 0.8020, -0.5866, -0.1129, and for BLUE: 0.8166, -0.5676, -0.1043. These coefficients define the chromatic mechanisms that are silenced by the respective unique hue settings. It is worth noting that the coefficients for RED and GREEN are very different, hence confirming the finding that RED and GREEN cannot be produced by the *same* linear chromatic mechanism (cf Figure 1: red and green are not collinear). The coefficients for YELLOW and BLUE on the other hand are very similar, which is consistent with the idea that the *same* linear mechanism underlies these unique hues (cf Figure 1: yellow and blue are collinear).

After having derived the best-fitting linear model, we now want to test how well it fits the data; we assume that each unique hue may be generated by a different chromatic mechanism (see above) as defined in Eq. 1-4. We evaluated the goodness of fit by calculating the perceptual error between the predicted and the observed unique hue settings for each of the four unique hues. The predicted unique-hue co-ordinates were obtained in the following way: for each (observed) unique-hue co-ordinate (in XYZ) we find the point on the theoretical plane which is closest (shortest Euclidean distance in XYZ space) to it. This is the predicted co-ordinate (in XYZ). We then converted the XYZ coordinates of both the observed settings and the predictions into the more uniform L*a*b* Space [11]. In figure 3, the observed (large circles) and predicted (small diamonds) settings are shown. The predictions were made by fitting the plane for each observer individually; for each unique hue we therefore estimated the coefficients (Eq. 1-4) for each observer separately. We then calculated the perceptual errors between the observed and the predicted unique hue co-ordinates using the colour difference formula proposed by the CIE in 1994, since this colour difference formula was developed to fit small colour differences [12, 13]. It might be worthwhile to point out that, in our calculations, this colour difference formula gave slightly lower values than the simple Euclidean distance in L*a*b* space. A perceptual error of 5 is visible in side-by-side image comparisons [14]. Figure 4 shows the distribution of these perceptual errors for all four unique hues. The black (left) bars indicate the relative frequency as a function of the error magnitude when the linear model was fitted for each observer separately ('INDIVID'). At least 80% of the errors are below 2 and less than 2% are above 5. The average error is less than 1.5 (see Table 1, row 2). This suggests that the linear model fits the data quite well.

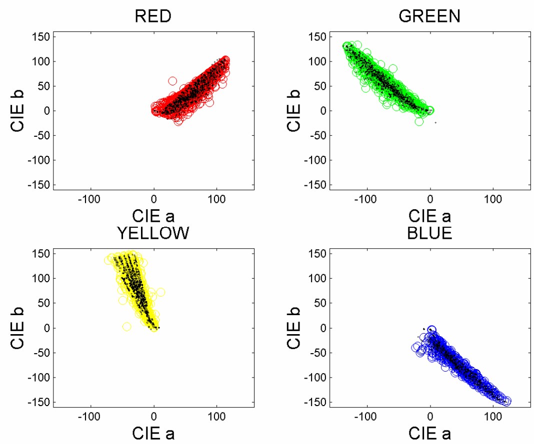

***Figure 3**. The unique hue settings are shown in the approximately uniform CIE L*a*b* diagram. Data are indicated by large circles; predictions are shown as small black diamonds. To visual the goodness of fit the predictions are plotted on top of the observed unique hue settings. The predictions shown here were made for each observer individually. We also fitted the pooled data and compared them to the individually derived predictions (see text for details).*

Table 1: MEAN PERCEPTUAL ERRORS (CIE 1994)

	RED	GREEN	YELLOW	BLUE
POOLED	2.5	1.8	2.1	1.5
INDIVID	1.4	1.1	1.1	1.3

Inter-observer Variability

The second aim of this study was to evaluate the variability between (colour-normal) observers in the perception of these unique hues. To this end we fitted the linear model to the pooled data (over all 18 observers) and made predictions for all observers using the same set of coefficients (one set of coefficients for each unique hue; cf Eq. 1-4). If there is little or no variability between observers, then the individually derived predictions (based on different coefficients for each observer) and the predictions derived from the pooled data should be very similar. We again calculated the perceptual errors (CIE 1994; see last section), i.e. the deviations of the observed unique-hue settings and the predictions, in Lab Space. The distribution of the errors is indicated by the white (right) bars in figure 4 (labeled 'POOLED). As expected, the fit of the linear model is worse for the pooled data compared to the individually fitted data. Predictions for the pooled data were arrived fitting only three free parameters (α, β, γ) for each hue; for the individual fits, 54 (3x18) parameters were fitted; for the 'pooled' data, 1616 data points were predicted with only three free parameters. Still, about 60% of the errors are below 2 and less than 5% are above a CIE L*a*b* error of 5. Similarly, the average error (Table 1, row 1) ranges from 1.5 to 2.5 and is higher (by a factor of about 1.5) than for the individually fitted data but still in an acceptable range [14] We therefore argue that the fit based on the pooled data is still acceptable and indicates a relatively small variability across observers with respect to these colour appearance mechanisms.

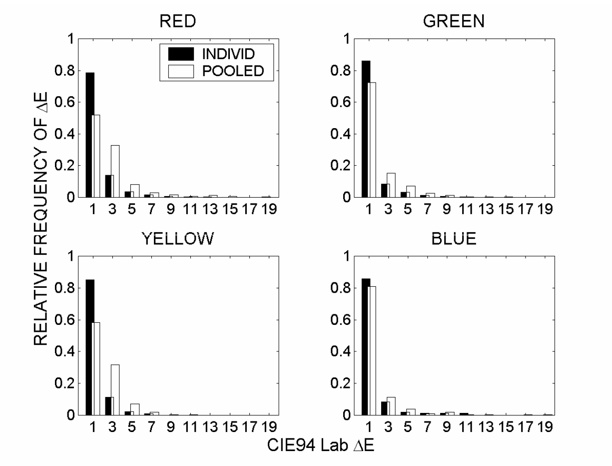

***Figure 4**. Histograms showing the distribution of the perceptual errors. The total number of trials for each colour was 1616. The perceptual errors are calculated by taking the difference between the predicted unique hue settings (based on the linear model, Eq. 1-4) and the observed unique hue settings (cf Figure 1) in L*a*b* space. Each panel shows the errors for a particular colour (RED, GREEN, YELLOW, BLUE). Within each panel two error distributions are shown: The left (black) bars shows the error frequencies when the predictions are made for each observer individually ('INDIVID'); the right (white) bars indicate the error frequencies when the data for all observers are fitted with a single set of parameters ('POOLED'). The discrepancy between the two error distributions reflects the inter-observer variability. For the individual fits (black bars) 80% of the errors are below 2. As expected, the goodness of fit of the linear model is slightly worse for the pooled fits, since only one set of parameters is estimated for all 18 observers (cf. Eq. 1-4); about 60% of the errors are below 2.*

Due to the relatively small inter-observer variability these unique-hue judgments might be useful to build standard models of colour appearance. It is likely that these unique-hue judgments are mediated by mechanisms that differ from mechanisms used for the discrimination of chromatic signals. This is corroborated by the finding that chromatic sensitivity and variability in unique-hue judgments is not correlated [5, 15] and by a recent study showing that the relative number of L :M cones does not affect unique yellow settings. [16, 17]. It is conceivable that relative L:M cone numbers affect chromatic discrimination performance, but have little effect on the chromatic appearance. This might be a factor contributing to the relatively small inter-observer variability.

Based on our perceptual error analysis, we suggest that the mechanisms that underlie these unique-hue judgments combine the XYZ co-ordinates linearly. This seems to be inconsistent with current successful colour appearance models [18-24] which usually include some non-linearity before the colour-opponent stage. We can think of two reasons for this discrepancy. The main reason is probably that we tested linearity in a rather restricted range determined by the gamut of our monitor. It is likely that non-linearities would have been revealed for higher intensities and more saturated colours. Secondly, the above argument that we might tap into different mechanisms depending on the judgment required of the observer, could also explain why we find no obvious deviations from linearity. The unique hues derived from a hue cancellation (or selection) task, might be generated by genuinely linear chromatic mechanisms.

Conclusion

We obtained unique hue settings from 18 observers using a visual display device. We conclude that, within the gamut of typical VDU, the higher-order colour mechanisms that yield the unique hues are linear mechanisms in XYZ space. Furthermore, the variability between observers is relatively small. These colour judgments may therefore useful to develop a standard model for colour appearance.

References

[1] A. Valberg, "A method for the precise determination of achromatic colours including white," Vision Research, 11, 157-160, (1971).

[2] B. A. C. Saunders and J. van Brakel, "Are there nontrivial constraints on colour categorization?" Behavioral and Brain Sciences, 20, 167-&, (1997).

[3] R. G. Kuehni, "Determination of unique hues using munsell color chips," Color Research and Application, 26, 61-66, (2001).

[4] B. E. Schefrin and J. S. Werner, "Loci of Spectral Unique Hues Throughout the Life-Span," Journal of the Optical Society of America a-Optics Image Science and Vision, 7, 305-311, (1990).

[5] M. A. Webster, E. Miyahara, G. Malkoc, and V. E. Raker, "Variations in normal color vision. II. Unique hues," Journal of the Optical Society of America A, 17, 1545-1555, (2000).

[6] D. Travis, *Effective Color Displays*. London: Academic Press, 1991.

[7] L. M. Hurvich and D. Jameson, "An opponent-process theory of color vision," Psychological Review, 64, 384-404, (1957).

[8] J. Larimer, D. Krantz, and C. Cicerone, "Oponenent-process additivity. I: Red/green equilibria," Vision Research, 14, 1127-1140, (1974).

[9] J. Larimer, D. Krantz, and C. Cicerone, "Oponenent-process additivity. II: Yellow/blue equilibria and nonlinear models," Vision Research, 15, 723-731, (1975).

[10] J. D. Jobson, *Applied Multivariate Data Analysis, volumes 1 and 2*. New York.: Springer-Verlag, 1991.

[11] G. Wyszecki and W. S. Stiles, *Color Science: concepts and methods, quantitative data and formulae*. New York: John Wiley & Sons, 1982.

[12] CIE, "Industrial color-difference evaluation, CIE publ. No. 116," Central Bureau of the CIE, Vienna, Austria,(1995).

[13] S. Y. Zhu, M. R. Luo, and G. H. Cui, "New experimental data for investigating uniform colour space," presented at CIE Expert Symposium, Scotsdale, Arizona, (2000).

[14] P. G. Engeldrum and J. L. Ingraham, "Analysis of white point and phosphor set differences of CRT displays," Color Research and Application, 15, 151-155, (1990).

[15] M. A. Webster, E. Miyahara, G. Malkoc, and V. E. Raker, "Variations in normal color vision. I. Cone-opponent axes," Journal of the Optical Society of America A, 17, 1535-1544, (2000).

[16] Y. Yamauchi, D. R. Williams, D. H. Brainard, A. Roorda, J. Carroll, M. Neitz, J. Neitz, J. B. Calderone, and G. H. Jacobs, "What determines unique yellow, L/M cone ratio or visual experience?" presented at SPIE, San Jose, (2003).

[17] J. Neitz, J. Carroll, Y. Yamauchi, M. Neitz, and D. R. Williams, "Color perception is mediated by a plastic neural mechanism that is adjustable in adults," Neuron, 35, 783-792, (2002).
[18] M. D. Fairchild and G. M. Johnson, "Color-appearance reproduction: Visual data and predictive modeling," Color Research and Application, 24, 121-131, (1999).
[19] M. D. Fairchild, "A revision of CIECAM97s for practical applications," Color Research and Application, 26, 418-427, (2001).
[20] M. D. Fairchild, *Color Appearance Models*. Chichester: Wiley-IS&T, 2004.
[21] N. Moroney, M. D. Fairchild, R. W. G. Hunt, C. Li, M. R. Luo, and W. T. Newsome, "The CAM02 Color Appearance model.," Proc. IS&T/SID 10th Color Imaging Conf., 23-28, (2002).
[22] M. D. Fairchild and G. M. Johnson, "iCAM framework for image appearance, differences, and quality," Journal Of Electronic Imaging, 13, 126-138, (2004).
[23] M. H. Brill, "Irregularity in CIECAM02 and its avoidance," Color Research And Application, 31, 142-145, (2006).
[24] C. J. Li and M. R. Luo, "Testing the robustness of CIECAM02," Color Research And Application, 30, 99-106, (2005).

Author Biography

Sophie Wuerger received her Ph.D. in Experimental Psychology at New York University (1991). Since then she has worked on visual perception, in particular colour vision, and is currently at the Centre of Cognitive Neuroscience, School of Psychology, University of Liverpool. She is a member of the Colour Group UK and the Optical Society of America.

.