Symmetries and asymmetries in chromatic discrimination

M. V. Danilova* and J. D. Mollon

1Laboratory of Visual Physiology, I. P. Pavlov Institute of Physiology, Nab. Makarova 6, St. Petersburg 199034, Russia
2Department of Experimental Psychology, Downing St., Cambridge CB2 3EB, UK
*Corresponding author: mvd1000@cam.ac.uk

Received October 3, 2013; revised December 24, 2013; accepted December 24, 2013; posted January 6, 2014 (Doc. ID 198671); published February 12, 2014

Under conditions of adaptation to a steady neutral field (metameric to Daylight Illuminant D65), forced-choice thresholds for color discrimination were measured for brief targets presented to the human fovea. Measurements were made along +45° and −45° lines in a MacLeod–Boynton chromaticity space scaled so that the locus of unique yellow and unique blue lay at −45°. The lines were symmetrical relative to the tritan line passing through the chromaticity of D65. Thresholds increased with distance of the probe chromaticity from D65. Thresholds were higher for saturation discrimination than for hue discrimination. A region of enhanced discrimination was found for thresholds measured orthogonally to the locus of unique blue and unique yellow. There may be an analogous enhancement near the loci of unique red and unique green. © 2014 Optical Society of America

OCIS codes: (330.1690) Color; (330.1720) Vision vision; (330.5020) Perception psychology; (330.5510) Psychophysics; (330.7320) Vision adaptation.

http://dx.doi.org/10.1364/JOSAA.31.00A247

1. INTRODUCTION

The color perception of the normal human observer depends on a comparison of the rates of quantum catch in three classes of cone, and therefore all physical colors—all spectral power distributions—can be represented as two ratios. In the familiar MacLeod–Boynton chromaticity diagram, for example, the x axis represents L/(L + M) and the y axis S/(L + M), where L, M, and S correspond to the quantum catches of the long-, middle-, and short-wavelength cones, respectively [1]; see Fig. 1.

When an observer is asked to discriminate two colors, independently of their luminance, his or she must ultimately depend on the change in the ratios of cone excitation from one stimulus to the other. However, discrimination thresholds vary nonuniformly across the chromaticity diagram and they critically depend on the chromaticity to which the observer is adapted [2]. Nor do subjective hue categories map in a simple way on to the MacLeod–Boynton diagram. In Fig. 1 we show the loci of “unique” hues, those colors that most observers judge to appear phenomenologically un mixed: blue, yellow, red, and green—plus white itself. An oblique line, running approximately from 475 to 575 nm, divides the diagram into reddish and greenish regions, and comprises lights that are pure blue, pure yellow, or white [5,6,9]. The boundary between lights that are bluish or yellowish is much less linear: the locus of unique reds runs nearly horizontally in the MacLeod–Boynton diagram but forms a firm angle with the locus of unique greens [5,9].

Is there any relationship between the subjective discontinuities in hue within the chromaticity diagram and the nonuniformities in difference limens for color? In the case of speech perception, optimal discrimination was classically found at category boundaries between consonants, e.g., between b and g [10]. Might the same be true for color? In two recent studies of color discrimination, one for parafoveal stimuli and one for foveal stimuli, we found a region of low thresholds close to the subjective category boundary between reddish and greenish hues, i.e., close to the locus of pure blues, pure yellows, and white [11,12]. These results are illustrated in Fig. 2. Figure 2A shows foveal thresholds for an individual observer, plotted directly in a portion of a MacLeod–Boynton diagram: thresholds were measured along lines approximately orthogonal to the subjective hue boundary, and each pair of data points represents the separation of chromaticities needed to sustain a criterion level of discrimination. In Fig. 2B are shown (i) the average settings of unique hues for five observers and (ii) the average positions of the minimal thresholds for the same observers. There is a fair coincidence between the subjective and the performance measurements. In other words, good discrimination occurs near the boundary between perceptual categories. To explain these results, one possibility would be to follow long tradition [13] and to postulate a chromatic channel that extracts a signal of the form (S + L)/M and gives rise to the sensations of redness and greenness. The equilibrium state of such a channel would correspond to the subjective category boundary, and we might expect differential thresholds to be lowest at the equilibrium state, since neural channels typically have a compressive, negatively accelerated response function [14].

Yet how firmly do the results of Fig. 2B require us to postulate a channel that extracts the ratio (S + L)/M and signals redness and greenness? Could the results be accounted for by the following rule? Thresholds are lower the closer the discriminanda are to the chromaticity of Illuminant D65. Local inspection of Fig. 2A suggests that this simple rule is inadequate [whether “distance” is measured in terms of L/(L + M) or S/(L + M) or both]. However, a clear test would...
be to perform a mirror-image version of the earlier experiment, reflecting the stimuli symmetrically around a vertical (tritan) line that runs through the white point in the MacLeod–Boynton diagram. Along a line that is a mirror image of the yellow–blue line, there is no equivalent category boundary. The locus of unique red runs more horizontally. Would there nevertheless be a minimum of thresholds near the line that is the mirror image of the yellow–blue line (line c in Fig. 3)? This is the test that we make in the present experiment.

Our experiment has eight conditions. In four conditions, thresholds are measured (as in Fig. 2A) along /0.0135 45° lines in a MacLeod–Boynton diagram, the ordinate of which has been scaled so that the yellow–blue line lies at −45°. In the other four conditions, thresholds are measured along −45° lines that are mirror images of the first set. Figure 3 shows the eight lines and their designations (A B C D for the +45° lines, a b c d for the −45° lines). The points along each line indicate the “referent” chromaticities at which we measured the difference limens. An additional advantage of the arrangement of Fig. 3 is that there are 16 referent chromaticities at which thresholds are measured in +45° and −45° directions around the same referent point: here the L/(L + M) and the S/(L + M) signals are each being modulated identically in the two conditions but in opposite phase relationships. Moreover, for the two lines passing through D65 we are able to compare thresholds that are saturation thresholds (those radial to D65) and thresholds that are hue thresholds (those in directions orthogonal to the radial line)—a distinction that proves to be critical.
2. METHODS

A. Apparatus and Stimuli

Measurements were made in Cambridge, England, and in St. Petersburg, Russia. In both laboratories, the stimuli were presented on calibrated Mitsubishi color monitors (Diamond Pro 2070) controlled by Cambridge Research Systems (CRS) graphics controllers (VSG 2/3 in Cambridge, Visage in St. Petersburg). In St. Petersburg, the monitor was set to a refresh rate of 92 Hz and a resolution of 1280 × 980 pixels; in Cambridge, the corresponding values were 100 Hz and 1024 × 768 pixels. The VSG system allowed outputs to be specified with a precision of 15 bits per gun, and the Visage, 14 bits. In both laboratories, the spectral power distributions of the monitor’s guns were measured with a JETI spectroradiometer, and the screens were linearized using a photodiode device (CRS “ColorCal” in Cambridge; “OptiCal” in St Petersburg). The algorithms for generating colors on the CRT screen were identical in the two laboratories.

The target field subtended 3° of visual angle and was divided into four sectors, one of which differed in chromaticity from the other three (see inset Figure 3). It was presented for 150 ms on a steady white background that had the chromaticity of CIE Illuminant D65 [16] and a luminance of 10 cd.m⁻². The sectors of the target array were separated by thin lines that had the chromaticity and luminance of the background. Fixation was guided by a diamond array of dark dots surrounding the area in which the target was presented. The display was viewed binocularly from a distance of 57 cm.

Chromaticities were specified in a MacLeod–Boynton diagram constructed from the cone sensitivities of DeMarco et al. [16]. The diagram represents a plane of equal luminance for the Judd 1951 Observer, where luminance is equal to the sum of the L- and M-cone signals [17]. The scale of the vertical ordinate of a MacLeod–Boynton diagram is arbitrary: we scaled our diagram so that a line running through 574 nm and the chromaticity of Illuminant D65 lay at −45°. Under the conditions of our experiments, this “yellow–blue line” represents the set of colors that are neither reddish nor greenish (as empirically measured in our earlier studies [12]).

The target field had an average luminance that was 30% greater than that of the background when expressed in the L+M units of our space; but to ensure that the observers could not discriminate the sectors on the basis of differences in sensation luminance, we jittered independently the L+M value of each sector by ±1% (in steps of 0.2%).

Although our targets were designed to remain approximately constant in photopic luminance, they necessarily varied in scotopic luminance. Could the results be affected by rod intrusion, despite the fact that the targets were brief, small, and centrally presented? By multiplying the actual spectral power distributions of the discriminanda by the scotopic luminosity function, we calculated that the maximal modulations of the rod signals at threshold were 5%–7% and most were much smaller, whereas the Weber fraction for rods is of the order of 30% [18]. Rod intrusion is therefore unlikely.

B. Procedure

Observers were asked to indicate by pushbuttons which quadrant of the target differed in chromaticity from the other three. Auditory feedback was given after each response. In any one experimental run, discrimination was measured along one of the eight lines of Fig. 3. Experimental runs were grouped into sets of eight, and within one set, the eight lines were tested in random order. There were six sets of experimental runs, the first set being treated as practice and not included in the analysis. Thus any given threshold for a given subject is based on five independent repetitions.

At the beginning of each experimental run, observers adapted to the neutral background field for 1 min before beginning measurements. Within each experimental run, thresholds were measured at a number of reference chromaticities. For lines A B C and a b c, there were six referents, and, owing to gamut limitations, for lines D and d there were five. These reference chromaticities were never themselves presented, but the chromaticities to be discriminated (the discriminanda) lay on the same line, straddling the reference point. Any one quadrant of the target could be the discrepant quadrant, and it was selected randomly. The discrepant quadrant could differ from the referent in either direction, and the remaining three quadrants then differed in the opposite direction. The chromatic separation of the discriminanda was increased or decreased symmetrically around the reference chromaticity according to the observer’s accuracy. The staircase procedure tracked 79.4% correct [19], and the separation between each of the discriminanda and the referent was adjusted in logarithmically equal steps. The reference and test chromaticities were expressed in terms of the abscissa of the MacLeod–Boynton diagram (i.e., the L/(L+M) or l, coordinates). At any one point on the staircase, one of the discriminanda had an l coordinate l₁, and the other had an l coordinate l₂, where l₁ was equivalent to the reference coordinate lᵣ multiplied by a factor a and l₂ was equivalent to lᵣ divided by a, where a is always >1.0. After three correct responses,
the value (α−1) was reduced by 10%, and after each incorrect response it was increased by 10%. The staircase terminated after 15 reversals, the last 10 reversal points being averaged to give the threshold. Within one experimental session, the reference stimuli were tested in random order.

C. Observers
All five observers had normal color vision as tested by the Cambridge Colour Test [20,21]. Observers 1 and 2 were the authors JM and MD, respectively. The other observers were highly practiced, but were naive as to the purpose of the measurements. Observers 2, 4, and 5 are female. All observers except observers 2 and 4 were tested in Cambridge. The experiments in both Cambridge and St. Petersburg were approved by the Psychology Research Ethics Committee of the University of Cambridge.

3. RESULTS AND DISCUSSION
In Fig. 4 we show the average results for our five observers. Discrimination thresholds are represented as dashes directly in the MacLeod–Boynton diagram, in the way that Wright [22] classically used dashes to represent thresholds in the CIE x, y chromaticity diagram. The dashes in Fig. 4 correspond to twice the separation of the discrimina that is needed to sustain a correct discrimination rate of 79.4%.

A. Distance from the White Point
Our observers were adapted to a steady neutral field metamerically to CIE Illuminant D65. A very clear result apparent in Fig. 4 is that thresholds increase systematically with the distance of the referent chromaticity from the chromaticity of this adapting field. Consider, for example, the uppermost pair of dashes in Fig. 4 (at the intersection of lines A and a).

Here the thresholds are more than 2.5 times those at the chromaticity of D65, even though the referent has the same L/(L + M) coordinate as D65. Only the level of S excitation distinguishes the two discrimination tasks.

In Fig. 5 we plot the 46 average thresholds directly against the Euclidean distance of the referent stimulus from the chromaticity of D65 in our space. Since the scaling of the ordinate of the MacLeod–Boynton space is arbitrary, the exact value of the Euclidean distance is specific to the scaling used in the present experiment. Nevertheless, from Fig. 5 it is clear that there is a very strong relationship between the magnitude of the threshold and the distance of the referent from the chromaticity of D65. The Spearman correlation coefficient is 0.753 and is highly significant (p < 0.001). (We give the rank-order correlation to emphasize that our conclusion is qualitative.)

Optimal discrimination at the adapting chromaticity is a familiar finding in studies of chromatic discrimination [8,23–25]. But the effect may be especially strong when there is little opportunity for adaptation to the target chromaticity during the threshold measurement [8]. In the present experiment, sensitivity is very briefly probed by our 150 ms targets, and the adaptive state of the eye is likely to be dominated by the D65 background field.

The strong increase of chromatic thresholds with distance from the adapting chromaticity may be seen as an analogue of Craik’s [26] finding for instantaneous luminance discrimination: although the human eye can achieve a Weber fraction of <1% for luminance over many log units of background luminance, this sensitivity is possible only if the observer is allowed to adapt to each new background level. In any one state of adaptation, there is a limited dynamic range of discrimination, centered on the background level. This result is usually

![Fig. 4. Average thresholds shown directly as dashes in a section of a MacLeod-Boynton chromaticity diagram. The dashes represent the separation of the discrimina needed to sustain a threshold performance of 79.4%. For visibility in the figure, the measured values have been doubled. The identification of the +45° and −45° lines is as in Fig. 3.](image-url)

![Fig. 5. Relationship between discrimination threshold and the Euclidean distance of the referent stimulus from D65. Thresholds are expressed as the factor by which the discrimina must differ from the reference chromaticity to sustain the criterion level of correct responses. Distance is expressed in terms of the scaling adopted in this paper for the S/(L + M) axis of the MacLeod-Boynton diagram. Notice the strong overall increase in thresholds according to the distance from the chromaticity with which the observer is adapted. In addition, the data points have been coded in terms of the angle formed between (i) a radial line from D65 to the referent chromaticity and (ii) the direction in which the threshold was measured. Cases where the angle is <45° are represented by open circles, and cases where the angle is ≥45° are represented as open triangles. The former can be regarded as predominantly measures of saturation thresholds, and the latter as predominantly measures of hue thresholds (see text).](image-url)
thought to reflect a shift in the equilibrium points of the neural channels used for the discrimination \[14\]. Recording from individual cells in the lateral geniculate nucleus of macaques, De Valois \textit{et al.} \[27\] observed behavior of this kind for chromatically opponent units. Psychophysical results from increment threshold measurements are also consistent with the principle that chromatically opponent channels have compressive response functions and have only a limited range of good discrimination \[28–30\].

If an observer is allowed to adapt to the particular color being tested, then good discrimination may be found over a larger range of the chromaticity diagram \[8\]. In classical experiments on color discrimination, such as those of MacAdam \[31\] and Wright \[22\], the adaptive state is likely to be different when measurements were made at different loci in the CIE chromaticity diagram. In MacAdam’s experiments, a large neutral surround was present, but the observer would necessarily become adapted to the chromatic region of the steady stimulus that was being used in the current match. MacAdam’s celebrated “discrimination ellipses” have sometimes been used to estimate the total number of colors that can be distinguished. We agree with Krauskopf and Gegenfurtner \[8\] that these estimates are misleadingly high: data such as those of Fig. \[4\] suggest that the human observer is likely to distinguish relatively few hues in any given state of adaptation.

**B. Discrimination as a Function of S-Cone Excitation**

Our stimulus matrix (Fig. \[3\]) was constructed so that we could measure thresholds at a number of referent stimuli that had the same \(L/(L + M)\) coordinate but different \(S\) coordinates. It is apparent from Fig. \[4\] that thresholds in our experiments depend on the level of S-cone excitation as well as on the ratio of \(L\) and \(M\)-cone excitation. This is made explicit in Fig. \[6\], where threshold factors are plotted for lines \(B\) and \(b\). Thresholds for referents with the same \(L/(L + M)\) coordinate plot vertically above one another in this figure. Such pairs of referents differ only in the level of \(S\)-cone excitation—and in the phase relationship of the \(S\) signal and the \(L/(L + M)\) signal. Clearly the level of \(S\) excitation, and its sign, have a large effect on the thresholds, and our results cannot be interpreted solely in terms of a chromatic channel that extracts the ratio of \(L\) and \(M\)-cone signals.

**C. Saturation Discrimination versus Hue Discrimination**

A general trend apparent in Fig. \[4\] is that discriminations in directions approximately radial to D65 are poorer than discriminations in tangential directions. In traditional terms, this is the distinction between saturation discrimination and hue discrimination. The \(+45°\) and \(-45°\) lines passing through D65 (lines \(C\) and \(c\)) in Fig. \[4\] offer the purest comparison of thresholds for saturation and hue: thresholds measured along each line correspond to saturation discrimination and orthogonal thresholds correspond to hue discrimination. In an ANOVA we formally compared saturation and hue thresholds for referents that lie along lines \(C\) and \(c\). The factors were observer, type of threshold (saturation versus hue), direction of line (\(+45°\) or \(-45°\)), and distance from the white point (two levels \(1, 2\)). There were significant effects of observer \((F[4] = 7.9, p = 0.001)\), type of threshold \((F[1] = 101.7, p < 0.001)\), and distance from the white point \((F[1] = 168.8, p < 0.001)\), but not of direction of line. There was, however, a significant interaction between direction of line and type of threshold \((F[1, 1] = 24.4, p < 0.001)\), and also between distance from the white point and type of threshold \((F[1, 1] = 16.8, p = 0.001)\).

For the pairs of “saturation” (radial) and “hue” (tangential) thresholds on lines \(C\) and \(c\), the saturation threshold for a given referent is on average 140% of the hue threshold. For any one of these pairs, the reciprocal modulation of \(L\) and \(M\) signals is around the same referent value but is combined in opposite phases with the \(S\) signal. The phase with which the two cardinal directions are combined has clear effects on the measured thresholds. So it is impossible to conclude that the cardinal axes act independently in chromatic discrimination. Previous results suggest the same conclusion \[32\].

How systematic is the superiority of hue discrimination over saturation discrimination? Is it possibly a general law that hue discrimination is better than saturation discrimination at any point in the chromaticity diagram for a given state of adaptation? A preliminary, but suggestive, answer is given by Fig. \[5\]. Here we have coded our thresholds at every referent chromaticity in terms of the angle between (i) a radial line from D65 to the referent chromaticity and (ii) the direction along which the threshold is measured. In the figure the open circles represent angles of less than \(45°\) from the radial and the open triangles represent angles equal to or more than \(45°\) from the radial. The former points are closer to being saturation thresholds, and the latter closer to hue thresholds—at least in the scaling that we have adopted for the chromaticity diagram. It is clear from Fig. \[5\] that saturation thresholds tend to be higher than hue thresholds. A more systematic study of radial versus tangential thresholds is warranted.

The superiority of hue discrimination over saturation discrimination is curious, and its neural basis remains to be
determined. An ecological explanation might be offered by the fact that natural surfaces typically reflect to the eye a specular component (representing the illuminant) as well the body color that derives from selective absorption by pigment molecules within the material. This was first pointed out by Monge in 1789 [23,24], and in some modern theories of color constancy it has suggested a way in which the visual system might recover the chromaticity of the illuminant by chromatic triangulation [35]. A typical surface is represented not by a point in chromaticity space but by a distribution of chromaticities that lie along a line between the body color of the surface and the color of the illuminant. The chromaticity of any given point on the surface moves to and fro along this line as the angle between the observer and the object varies. The lines under discussion here are, of course, lines of varying saturation. In identifying objects, we may rely more on the direction of the line (i.e., the hue of the surface) than upon the exact position of the chromaticity along the line (i.e., the saturation), since the latter is more variable in normal experience.

D. Comparison of +45° and −45° Lines
A central question in our study was whether color discrimination thresholds are symmetrical for +45° and −45° lines. Our stimulus matrix (Fig. 3) was constructed so that each referent value on lines A B C D has an equivalent referent on mirror-image lines a b c d. We formally tested the question of symmetry with a two-way ANOVA with factors direction (−45° versus +45° lines) and referent (the 23 individual thresholds tested on each set of lines). Both direction and referent were highly significant (F[1] = 4.483, p < 0.001; F[22] = 38.8, p < 0.001), as was their interaction (F[1, 22] = 4.3, p < 0.001). We conclude that color discrimination along the lines a b c d is not symmetrical to that along the lines A B C D.

We also performed such an ANOVA for just the 16 referents in the matrix (Fig. 2), where thresholds are measured in the −45° and +45° directions at the same point in chromaticity space. The factors were again direction (−45° versus +45° lines) and referent (the 16 individual thresholds tested on each set of lines). Both factors were again highly significant, as was their interaction. In this case, for each referent—at each point in the chromaticity space—the analysis is comparing thresholds for modulations that differ only in the phase with which L/M and S are combined. Yet there are significant differences between thresholds measured in −45° and +45° directions—although it is not always the same phase of S and L/M that is favored.

E. Color Discrimination and Subjective Category Boundaries
Figure 7A shows thresholds measured along the four +45° lines A B C D plotted against the L/(L+M) coordinate of the referent stimulus. The ordinate represents the factor by which the discriminanda must each differ from the referent to sustain 79.4% correct performance. The data are averages for five observers, and the functions fitted to the data are inverse third-order polynomials. For each line, a vertical arrow in the diagram indicates the average positions of the stimulus that subjectively appeared neither reddish nor greenish in our previous measurements [12]. These results confirm our previous finding that forced-choice performance thresholds are minimal near the subjective category boundary.

Figure 7B shows the corresponding thresholds for the four −45° lines, a b c d. The results for line c, which passes through the chromaticity of D65, are almost symmetrical to those for the counterpart line, C, and in both cases, of course, discrimination is optimal at the L/(L+M) value of the adapting background—lower indeed than in any other conditions. Lines D and d, which intersect at a chromaticity below that of D65, also show a high degree of symmetry. However, lines a and b exhibit results that are less obviously symmetric to those
of their counterparts A and B; for example, thresholds on line a continue to fall to the highest L/(L + M) value tested. By vertical arrows we indicate the L/(L + M) values at which lines a, b, c, d cross the loci of unique red and unique green (cf. Fig. 1). We have taken these approximate values from the literature (e.g., [2]), and we did not obtain phenomenological estimates of unique red and unique green in the present experiment. However, there is a suggestive alignment of the threshold minima with the estimated hue boundaries, and this second coincidence of performance and phenomenal measures deserves further investigation.

4. CONCLUSIONS

1. If chromatic discrimination is probed without disturbing the state of adaptation, the single most important factor in determining thresholds is the distance, in the chromaticity diagram, of the probe from the adapting chromaticity.

2. Thresholds measured along radial 45° lines from the adapting chromaticity (i.e., saturation thresholds) are higher than thresholds measured orthogonally to these lines (i.e., hue thresholds). It may be a general rule that saturation thresholds are larger than hue thresholds.

3. Thresholds measured along 45° lines are not symmetrical to those measured along 45° lines.

4. Good discrimination is found in the vicinity of the boundary between reddish and greenish hues (i.e., when thresholds are measured orthogonally to the unique-yellow/unique-blue locus.)

5. There may also be a region of good discrimination in the vicinity of the boundary between yellowish and bluish hues (i.e., when thresholds are measured orthogonally to the unique-red/unique-green locus.)

ACKNOWLEDGMENTS

This work is supported by Russian Foundation for Basic Research grant 12-04-01797 and Royal Society International Exchanges grant IE110252. We thank Dr. M. Webster and Dr. J. Bosten for discussion.

REFERENCES


