COLORBLIND VISION

I. LUMINOSITY LOSSES IN THE SPECTRUM FOR DICHROMATS*

BY SELIG HECHT† AND YUN HSIA‡

(*From the Laboratory of Biophysics, Columbia University, New York*)

(Received for publication, June 18, 1947)

I

Nature of Work

In recent years the shape and location of the spectrum luminosity curves of colorblind individuals have been determined with good precision (Pitt, 1935; Hecht and Shlaer, 1936). Fig. 1 presents the average measurements of six deuteranopes and six protanopes, as found by Pitt. The measurements by Hecht and Shlaer on three subjects are in agreement with Pitt's data. Fig. 1 shows that compared to the normal luminosity, the protanope curve is displaced considerably toward the blue, whereas the deuteranope curve is only slightly displaced toward the red.

The measurements in Fig. 1 are drawn in the traditional way of considering each luminosity curve by itself and placing the point of maximum brightness for each at 100 per cent. The question arises as to whether the three curves really have the same height, and if not, what their relative heights are. This is particularly relevant to the problem of what colorblindness is and what relation it has to normal color vision.

Normal color vision is best described in terms of a three receptor system; and colorblindness is often assumed to represent the loss or inactivation of one of these receptor systems. If this is true, then certainly the loss of one of the three receptor systems should show itself in some loss of brightness in the spectrum as experienced by the colorblind. One might expect the colorblind curves in Fig. 1 to have lower maxima than the normal. Abney (1913) drew such lowered luminosity curves for colorblinds. However, these reduced curves were not found from measurements, but were invented in terms of theory (Abney, 1913, page 281).

Because of this total lack of knowledge, we decided to find the relative heights of the three curves in Fig. 1 by measuring the actual visual thresholds of color-

---

*A report on these findings was given to the Optical Society of America at its meeting on October 3 to 5, 1946 in New York and to the International Conference on Colour Vision in Cambridge, England, July 28 to 29, 1947.

†Died, September 18, 1947.

‡Fellow of the China Institute of Physiology and Psychology.

blinds and normals in different parts of the spectrum. The idea is to determine for a given portion of the spectrum the amount of light required at the foveal threshold of normal persons and of colorblind persons, and to do this for different parts of the spectrum so chosen as to cover the spectral range.

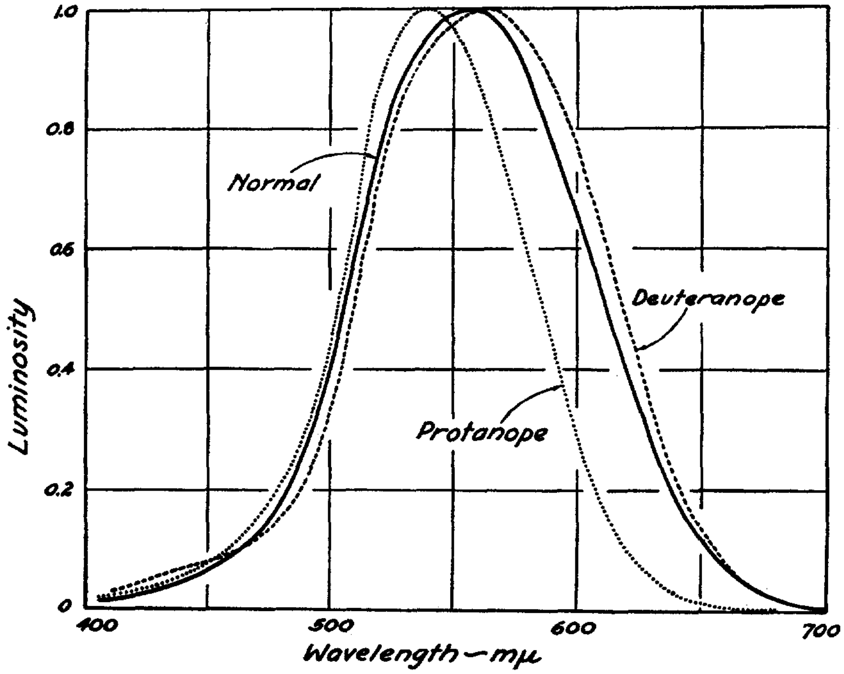

FIG. 1. Spectrum luminosity distributions for normals and colorblinds. The maximum for all curves is arbitrarily put at 100. Data from Pitt (1935).

II

Selection of Subjects

For our measurements we used nine color-normals, six protanopes, and seven deuteranopes, all men between 20 and 35 years of age.

All subjects were tested for their color vision in several ways. The first test consisted in reading the Ishihara pseudoisochromatic plates, "Tests for Colorblindness," 5th edition. Following this, the subject read Stilling's "Pseudo Iso-Chromatische Tafeln," 19th edition. Anyone who read correctly all the plates in these two tests without hesitation was tentatively considered normal. He was then further tested for the Rayleigh equation with an anomaloscope made in our own Laboratory. The equation requires the subject to match a yellow of 575 mμ with a mixture of 555 mμ green and 635 mμ red. A subject was considered normal only when his match was sharp and fell in the narrow range found by us in unpublished studies with many normal persons.

The colorblind subjects were all complete dichromats, not anomalous trichromats. We first chose those persons who made extensive errors in the Ishihara and Stilling tests. Of these we selected only those who in the anomaloscope test were able to match perfectly the yellow of 575 mμ with the green of 555 mμ by itself, and also with the red of 635 mμ by itself. By comparing the relative brightness at match of the yellow and red we classified the subject as protanope or deuteranope; if the red was much brighter (to us) than the yellow he was called protanope (or loosely, red-blind), whereas if at match the red and yellow were equally bright, he was called a deuteranope (or green-blind). The final critical test was always to determine the presence of a neutral point, or white spot, in the spectrum. Only those subjects who matched a sharply located neutral point in the spectrum with whole white light of 5000°K as a standard were considered as complete dichromats, and were used as subjects for the measurements. The neutral point determinations were made with a modified Helmholtz color mixer (Hecht and Shlaer, 1936).

TABLE I

Filters for Isolating Spectrum Regions

Spectral region	Filters	Central λ
		mμ
Blue	C 511 + C 368	469
Blue-green	W 75 + C 978	500
Green	W 74	535
Yellow	W 73 + C 978	579
Red	C 241 + C 978	651

C means Corning glass filter; W means Wratten gelatin-in-glass filter.

III

Apparatus and Procedure

The measurements of the foveal thresholds were made with the original adaptometer designed for just such purposes (Hecht and Shlaer, 1938). In this instrument the intensity of the test light is varied with neutral filters and a neutral wedge, while the color of the light is controlled by appropriate color filters.

Five regions of the spectrum were isolated by means of the filter combinations given in Table I. The transmissions of the individual filters were measured with a photoelectric spectrophotometer (Shlaer, 1938). The transmissions of the filter combinations are shown in Fig. 2. For each filter combination the transmissions of the neutral filters and of the wedge were measured with a Martens polarization photometer. The color temperature of the lamp of the adaptometer was measured with an Eastman color temperature meter and found to be 2700°K.

The absolute brightness of the test field in white light was measured directly at a chosen wedge setting with a Macbeth illuminometer placed at the exit pupil of the adaptometer. Knowing this brightness, and knowing the transmissions of the neutral and color filters, the slope of the wedge, and the color temperature of the lamp, one can compute precisely the spectral composition and the relative energy content of the test fields used in the measurements. Since these calibrations are merely detailed steps in the computation of the final results, they need not be given here.

Observations were monocular, and the subject fixated a tiny red point reduced to the lowest possible brightness by the subject himself. The circular test field was central, and was 1° in diameter. Exposures were in flashes of ⅕ of a second, and were controlled by the subject. The experimenter arranged the wedge and the filters, and the subject manipulated the shutter when he was ready to observe the flash. He merely reported the presence or absence of light.

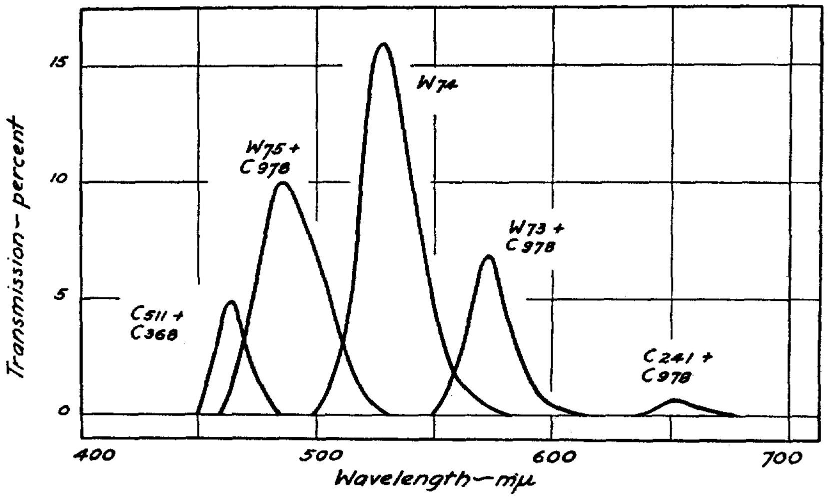

FIG. 2 Transmission of filter combinations for isolating parts of the spectrum. *C* is Corning and *W* is Wratten.

The natural pupil was used. Because the subjects were all dark-adapted, the pupil was maximal. Variations in the pupil size at maximal pupil opening are of only trifling influence on cone vision. This is due to the Stiles-Crawford phenomenon (Stiles and Crawford, 1933) which renders the light passed by the peripheral portion of the expanded pupil extremely inefficient in stimulating the cones.

Cone adaptation is usually over in 10 minutes of darkness, even after exposure to very bright light. We allowed our subjects at least 15 minutes in the dark after a preliminary stay in the laboratory before beginning observations so that cone dark adaptation was complete.

An experimental session consisted of determining three times in succession the final threshold for seeing the central 1° field at the five selected positions of the spectrum. We always started with the blue end of the spectrum and worked toward the red. Most of our subjects came for four such sessions. During the first session the measurements were recorded but not counted because we considered the period as one of instruction and practice. The remaining three separate sessions constituted the final measurements.

IV

Relative Luminosities

Our interest in the present study is to determine the differences in foveal threshold between colorblinds and normals. To establish the normal base-line, we have averaged the measurements for the nine color-normals at each spectrum locus, and have put the average threshold value equal to unity.[1] Table II then shows the logarithmic deviations of the nine observers from the average values and gives some idea of the individual variations encountered.

In the same Table II the thresholds of the protanopes and of the deuteranopes are also shown as deviations from the average color-normal. From these deviations one can assess the magnitude and the validity of the average differences between normal and colorblind persons.

It is apparent from Table II that beginning with the blue and going toward the red, the threshold of the protanope, compared to the normal, steadily rises in the spectrum. The two thresholds are practically identical in the blue, but the protanope threshold is well over 1 log unit greater than the normal in the red. The deuteranope threshold is also very nearly normal in the blue; but it definitely rises in the green and remains at about that level through the rest of the spectrum. The average measurements are shown in Fig. 3 where the normal is represented by a horizontal line and the thresholds of the two kinds of colorblinds are shown displaced above the normal the appropriate logarithmic distance.

In order to see what these measurements actually do to the shapes and positions of the colorblind luminosity curves, we need to place them in relation to the normal luminosity distribution in the spectrum. This is done in Fig. 4 with relation to the color-normal curve taken from Wald's recent study (Wald 1945). The reason for choosing Wald's data is that they were secured in essentially the same way as ours, that is, by measurements of the energy threshold of a 1° field in the central fovea after complete dark adaptation. However, with only minor differences Wald's normal curve is practically the same as the standard luminosity curve for the normal eye which has been used for years (Gibson and Tyndall, 1923).

In Fig. 4 at the selected five points on the spectrum, the average logarithmic

---

[1] The actual averages of the brightness thresholds for the nine normal subjects for the spectrum loci 469, 535, 579, and 651 mμ are respectively 6.13, 6.15, 6.13, and 6.17 in log micromicrolamberts. These threshold values are essentially identical, as they should be. However, the average threshold for 500 mμ for the same subjects is 6.32 in log micromicrolamberts, that is, about 0.15 log unit higher than it should be. Recalibration of all the filters and the wedge, and reworking of all the computations yielded no change in these values. Since we are interested only in differences between normal and colorblind observers this curious deviation has no relevance.

differences found in Table II are drawn for the protanope and deuteranope as displacements downwards from the normal curve of Wald. Through these displaced points we have drawn smooth curves. Fig. 4 is therefore the same as Fig. 3 except that instead of the normal curve being rectified, it is now presented in its real spectral form. Fig. 4 shows unequivocally that for both types of colorblindness there occurs not merely a shift of maximum or a change in shape of the curve, but a real loss of luminosity in the spectrum.

TABLE II

Foveal Thresholds of Colorblinds Compared to Normal for 1° Field

The average normal threshold is put at 1. All values are given as logarithmic deviations from the average normal (log 1 = 0).

Color vision	Subject	Deviation from normal mean in log units at				
		469 mμ	500 mμ	535 mμ	579 mμ	651 mμ
Normal	1	−0.04	−0.01	+0.06	+0.06	−0.03
	2	+0.06	+0.05	+0.02	+0.02	+0.15
	3	−0.04	−0.07	+0.01	+0.08	+0.07
	4	+0.10	+0.01	−0.09	−0.12	−0.09
	5	+0.05	+0.11	+0.08	+0.06	−0.03
	6	−0.03	−0.05	−0.14	−0.05	−0.01
	7	−0.12	−0.15	−0.11	−0.13	−0.10
	8	+0.01	+0.07	+0.07	+0.02	+0.04
	9	−0.01	+0.01	+0.08	+0.09	0.00
	Average.......	0	0	0	0	0
Protanope	1	+0.03	+0.10	+0.29	+0.63	+0.87
	2	−0.03	+0.01	+0.21	+0.49	+1.37
	3	+0.11	+0.11	+0.15	+0.29	+1.06
	4	+0.08	+0.14	+0.19	+0.23	+1.21
	5	+0.11	+0.14	+0.14	+0.20	+1.05
	6	+0.06	−0.05	+0.14	+0.15	+1.11
	Average.......	+0.06	+0.08	+0.19	+0.33	+1.11
Deuteranope	1	−0.12	−0.02	+0.09	+0.07	+0.04
	2	+0.19	+0.30	+0.30	+0.30	+0.34
	3	+0.13	+0.11	+0.33	+0.26	+0.21
	4	+0.16	+0.11	+0.19	+0.20	+0.23
	5	+0.08	+0.11	+0.29	+0.20	+0.03
	6	+0.01	+0.10	+0.21	+0.23	+0.21
	7	+0.01	+0.08	+0.22	+0.28	+0.26
	Average.......	+0.07	+0.11	+0.23	+0.22	+0.20

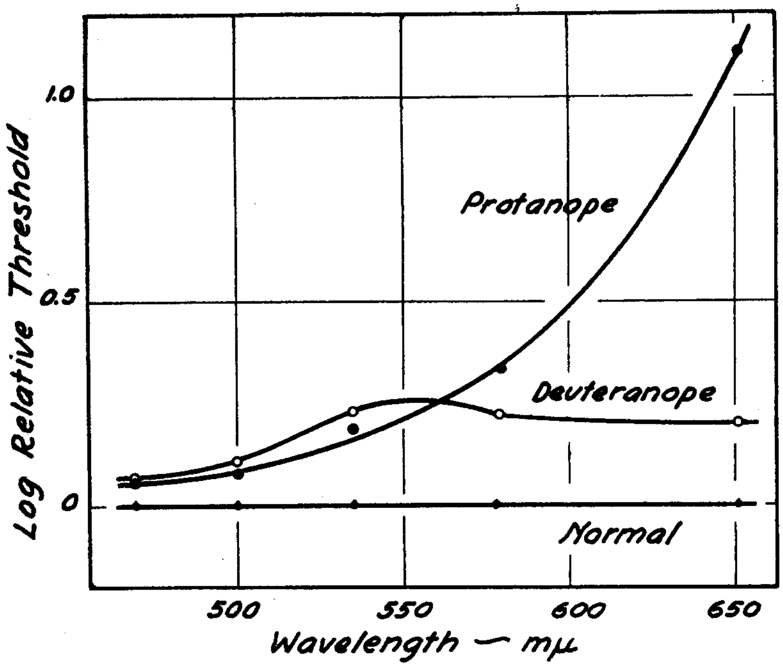

FIG. 3. Average foveal thresholds of six protanopes and seven deuteranopes compared to nine normals. Test field is circular and 1° in diameter.

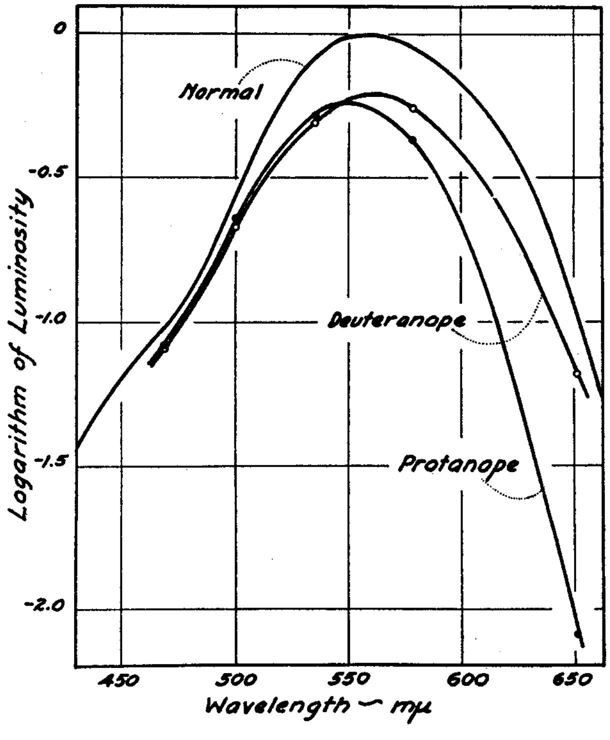

FIG. 4. Real luminosity distributions in the spectrum of normals and dichromats.

To determine the precise amount of this loss, we have plotted the data in arithmetical form in Fig. 5. The area under such an arithmetical luminosity curve represents the total brightness of an equal energy spectrum. We have measured with a planimeter the relative areas under the three curves. If the area under the normal curve is put equal to 100, the area under the protanope curve turns out to be 51.0, whereas the area under the deuteranope curve is 61.2.

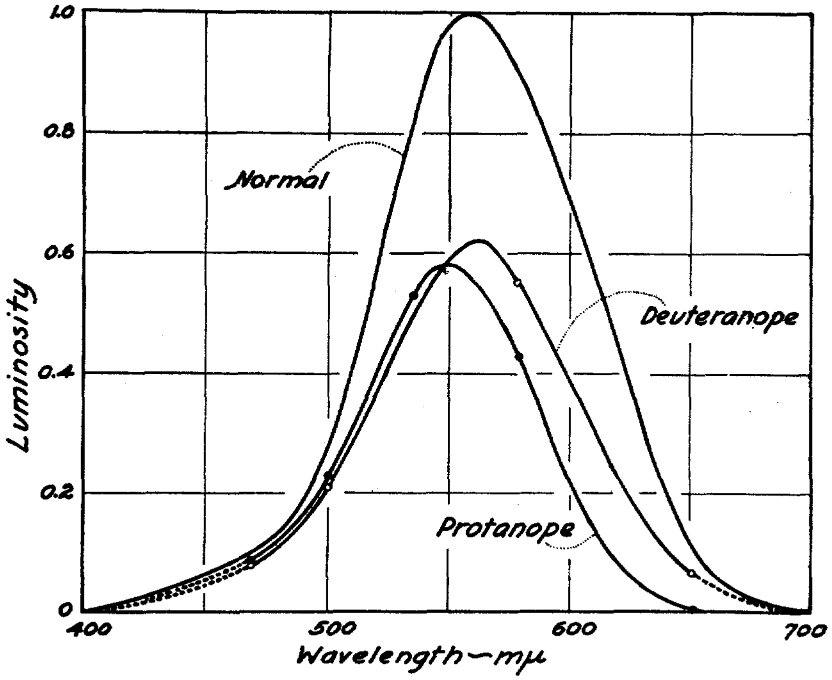

FIG. 5. Arithmetical luminosity distributions in the spectrum to show the real heights of the colorblind curves compared to the normal. The areas under the curves indicate the total brightness of an equal energy spectrum. When the normal area is 100, the protanope area becomes 51.0 and the deuteranope area becomes 61.2.

Evidently, compared to the normal, the protanope loses almost one-half the luminosity of the spectrum while the deuteranope loses almost two-fifths the luminosity of the spectrum. These losses are large, and must have meaning for color vision theory.

V

Color Vision and Colorblindness

There is no completely adequate theory of color vision. However, the one reasonable basis for such a theory has consistently been Young's notion that there are three receptor systems in the retina (Young, 1807) which may be designated as *B*, *G*, and *R* to indicate their essentially qualitative uniqueness in yielding respectively blue, green, and red sensations when brought into action by light. Each receptor system produces only the sensation unique for it, regardless of the part of the spectrum which sets it into action, and the sensations produced by various parts of the spectrum result from the combined action of these three systems in different degrees. Certain combinations produce specific effects. Thus the combined actions of the *G* and *R* systems result in the unique sensation of yellow, while the combined actions of *B*, *G*, and *R* result in the unique sensation of white (*cf.* Hecht, 1928). Moreover, the action of the receptors contributes brightness as well as color, and the brightness contributions of the three systems are strictly additive.

Young supposed that colorblindness is due to the loss of one of these three receptor systems. This supposition is still held in one form or another (Pitt, 1944).

Our present measurements support this simple and direct formulation for colorblindness. If a protanope has lost the *R* receptor system, he should lack not only its color effects but also its brightness contribution, and the loss should correspond to the contribution which the *R* receptor system makes to the normal luminosity of the spectrum. Since our data show that the protanope suffers a loss of 49 per cent in spectrum luminosity, it would seem that the *R* system normally contributes 49 per cent of the brightness of an equal energy spectrum. In the same terms, the loss of 39 per cent of spectrum luminosity by the deuteranope corresponds to his loss of the *G* receptor system, and would indicate that the *G* system contributes 39 per cent of the normal spectrum brightness.

Since the total brightness is supplied by $B + G + R$, and since $G + R$ together contribute 88 per cent, the remaining 12 per cent must represent the contribution of the *B* receptor system alone. It follows from this reasoning that if tritanopia—the third form of colorblindness—results from the loss of the *B* receptor system, it should involve the loss of only 12 per cent of the brightness of an equal energy spectrum. No measurements now exist to judge of this, and it therefore represents a specific prediction in terms of Young's idea of colorblindness.

The relative brightness contributions of the *B*, *G*, and *R* systems as 12, 39, and 49 per cent respectively are roughly the same as many previous estimates of the contributions of these three systems (Ives, 1923; Wright, 1929–30). In general the *G* and *R* systems have been given approximately equal weight in normal luminosity, whereas the *B* system has most often been evaluated as even less than 10 per cent. The present measurements do not support the notion (Hecht, 1931) that the three receptor systems contribute equally to the brightness of the spectrum.

Even though the simple idea that colorblindness is due to a loss of one receptor system accounts for our luminosity measurements, it does not account for other equally important aspects of colorblindness, particularly the color sensations. If in normal vision the varying degrees of activity of all three systems, *B*, *G*, and *R*, produce all the varied color sensations, then the loss of one system should reduce the number and alter the quality of these sensations. For example, if the *R* system is lost, no red sensations should be possible. Loss of the *R* system should leave the spectral gamut as made up only of blue and green sensations. Moreover, since yellow occurs from the combined activities of the *G* and *R* systems, the protanope should have no yellow sensation, and indeed no white sensation either because the action of all three systems is required for this effect.

A similar situation must obtain for the deuteranope also. Having lost the *G* system, he should see neither yellow nor white as we do. Moreover, the tritanope, because of his lost *B* system, should also be unable to see white as do normals; to him it should appear yellow because of the action of the *G* and *R* systems alone.

None of these consequences of the loss of a receptor system is true. Colorblind persons, beginning with Dalton (1798) who described his sensations with great clarity, insist that they see white as uniquely colorless. Their insistence is confirmed by those occasional individuals who are normal in one eye and colorblind in the other. The best described case is that of a tritanope (Dieter, 1927) who on comparing the spectrum and other colors with his two eyes described white as the same with both eyes. In particular, the neutral points at 575 mμ and 415 mμ were unequivocally described as white. Other instances of monocular colorblindness (von Kries, 1919) bear out these facts about white.

Moreover, in terms of the loss of a receptor system, neither the deuteranope nor the protanope can have the sensation of yellow. Yet these dichromats consistently describe the spectrum as made up of two hues: blue on the short-wave side of the white neutral point, and yellow on the long-wave side of the neutral point. In addition they record that these two hues are most saturated at the extremes of the spectrum, and gradually become unsaturated toward the neutral point which is completely unsaturated as white. In fact, colorblinds make wavelength discriminations in the spectrum on either side of the neutral point not in terms of hue but in terms of saturation (Hecht and Shlaer, 1936). These descriptions are corroborated by von Kries' monocular colorblind. In short the sensory reports given by dichromats cannot be accounted for by Young's idea of the simple loss of one receptor system.

The sensations reported by colorblinds may be accounted for by an alternate proposal first suggested by Fick (1879). According to this proposal, colorblindness involves the transformation of the spectral sensitivity of one receptor system into that of one of the other receptor systems. For example, suppose that the *R* system is changed so that its sensitivity in the spectrum becomes identical with the G system. The altered R receptor system and the normal G system will now be equally stimulated by light of any part of the spectrum. This applies only to the outermost light-receiving end of the receptor elements. The rest of the altered R system beyond the receptor elements will remain as before. Thus regardless of the light which it receives, it will still transmit impulses which will result in the production of the unique red sensation. Such an eye will not be able to discriminate hues in the spectrum on the long-wave side of the neutral point because this whole region will be yellow of different degrees of saturation.

This formulation of colorblindness accounts for the sensations which colorblinds have, and for the errors and confusions which they make. Unfortunately it is not supported by our present measurements. For instance, if the R system is not lost, but its sensibility distribution in the spectrum is merely altered, then there should be no loss in the total spectrum luminosity of protanopes. Yet our measurements show a 50 per cent loss. And a similar situation obtains for a deuteranope, whose loss is almost 40 per cent. Thus the idea of a simple loss of one receptor system accounts for the luminosity measurements but completely fails to account for the sensations; whereas the idea of a transformation of one receptor system into another accounts for the sensations, but cannot encompass the luminosity loss data.

Pitt's recent (1944) proposal that protanopia represents a simple loss of R whereas deuteranopia represents a transformation of G into R does not resolve this dilemma. Both types of colorblind show a loss of luminosity, and the sensations of the protanope still remain unaccounted for by a mere loss of receptor system.

We thus have two mutally exclusive suggestions for the basis of colorblindness each of which describes only one aspect of the phenomenon. Clearly, a new formulation for colorblindness is required which will combine the virtues of both ideas and eliminate their contradictions.

SUMMARY

1. Measurements have been made of the dark-adapted foveal threshold of normal and colorblind persons in five parts of the spectrum using a 1° circular test field.

2. Compared to normals, protanopes (red-blinds) show an elevation of the threshold which increases slowly from blue to yellow and rises rapidly thereafter until in the red the threshold is more than ten times as high as normal. Deuteranopes (green-blinds) do not show so high an elevation, their maximum in the green being only about 70 per cent above normal.

3. These threshold elevations correspond to luminosity losses in the spectrum. For the protanope the total loss in the spectrum is nearly one-half of the normal luminosity; for the deuteranope it is nearly two-fifths of normal.

4. Such losses support the idea that colorblindness corresponds to the loss of

one of the three receptor systems usually postulated to account for normal color vision. However, the color sensations reported by colorblind persons, especially monocular colorblinds, do not support the idea of a lost or inactivated receptor system. A fresh explanation for colorblindness is called for to reconcile these conflicting kinds of evidence.

BIBLIOGRAPHY

Abney, W. de W., Researches in colour vision and the trichromatic theory, London, Longmans Green and Co., 1913.

Dalton, J., Extraordinary facts relating to the vision of colours, *Mem. Lit. and Phil. Soc., Manchester*, 1798, **5,** 28.

Dieter, W., Über die subjektiven Farbenempfindungen bei angeborenen Störungen des Farbensinnes, *Z. Sinnesphysiol.*, 1927, **58,** 73.

Fick, A., Die Lehre von der Lichtempfindung, in Handbuch der Physiologie, (L. Hermann, editor), Leipzig, Vogel, 1879, **3,** pt. 1, 139.

Gibson, K. S. and Tyndall, E. P. T., The visibility of radiant energy, *Bureau Standards Scient. Papers*, 1923, **19,** 131.

Hecht, S., On the binocular fusion of colors and its relation to theories of color vision, *Proc. Nat. Acad. Sc.*, 1928, **14,** 237.

Hecht, S., The interrelation of various aspects of color vision, *J. Opt. Soc. America*, 1931, **21,** 615.

Hecht, S., and Shlaer, S., The color vision of dichromats. I. Wavelength discrimination, brightness distribution, and color mixture, *J. Gen. Physiol.*, 1936, **20,** 57.

Hecht, S., and Shlaer, S., The color vision of dichromats. II. Saturation as the basis for wavelength discrimination and color mixture, *J. Gen. Physiol.*, 1936, **20,** 83.

Hecht, S., and Shlaer, S., An adaptometer for measuring human dark adaptation, *J. Opt. Soc. America*, 1938, **28,** 269.

Ives, H. E., The transformation of color mixture equations from one system to another: II. Graphical aids, *J. Franklin Inst.*, 1923, **195,** 23.

von Kries, J., Über einen Fall von einseitiger angeborener Deuteranomalie (Grünschwäche), *Z. Sinnesphysiol.*, 1919, **50,** 137.

Pitt, F. H. G., Characteristics of dichromatic vision, with an appendix on anomalous trichromatic vision, *Great Britain Med. Research Council, Special Rep. Series, No. 200*, 1935.

Pitt, F. H. G., The nature of normal trichromatic and dichromatic vision, *Proc. Roy. Soc. London, Series B*, 1944, **132,** 101.

Shlaer, S., A photoelectric transmission spectrophotometer for the measurement of photosensitive solutions, *J. Opt. Soc. America*, 1938, **28,** 18.

Stiles, W. S., and Crawford, B. H., The luminous efficiency of rays entering the eye pupil at different points, *Proc. Roy. Soc. London, Series B*, 1933, **112,** 428.

Wald, G., Human vision in the spectrum, *Science*, 1945, **101,** 653.

Wright, W. D., A re-determination of the mixture curves of the spectrum, *Tr. Opt. Soc. London*, 1929-30, **31,** 201.

Young, T., On the theory of light and colors, in Lectures in Natural Philosophy, London, Joseph Johnson, 1807, **2,** 613.