DEFECTIVE COLOR VISION AND ITS INHERITANCE

By George Wald*

BIOLOGICAL LABORATORIES, HARVARD UNIVERSITY

We owe the first accurate description of color blindness to the chemist John Dalton, who in 1798 characterized his own condition in the words "persons in general distinguish six kinds of colour in the solar image; namely, red, orange, yellow, green, blue, and purple. . . . To me it is quite otherwise:—I see only two, or at most three, distinctions. These I should call yellow and blue; or yellow, blue, and purple. My yellow comprehends the red, orange, yellow, and green of others; and my blue and purple coincide with theirs. That part of the image which others call red, appears to me little more than a shade, or defect of light. . . ." In this superb paper, Dalton set all the main themes for the study of color blindness. He pointed out that his vision had always been this way; that his brother had the same defect; that he knew of nearly 20 such cases, all male; and that in all but perhaps one instance the parents and children of such persons had normal vision.[1]

Dalton thought the trouble to be that his vitreous humor was tinted blue, and so did not pass red light; but Thomas Young said of this: ". . . it is much more simple to suppose the absence or paralysis of those fibers of the retina which are calculated to perceive red."[2] Dalton held to his own view, and the argument was settled only by autopsy at his death in 1842, when his eyes were found to be normally transparent to red light.[3] Color blindness is still called *daltonism* throughout Europe; but Young's explanation, that it involves the loss of one of three normal color mechanisms, remained for many years the dominant theory of this condition.

This is, however, as Maxwell later made plain, more a mathematical than a physiological theory.[4, 5] It implies only that normal human color vision involves the operation of three independent variables, hence is *trichromatic*, with no indication of what those variables might be. The most obvious expression of this condition is that normal persons must mix three different colors in various proportions to match all the hues they see. Most color-blind persons are *dichromats:* their vision involves the play of only *two* variables, and they can match all discriminable hues by mixing two colors. There is also a rare class of *monochromats* who can match all colors with one another by adjusting the brightness. Finally there are the *anomalous trichromats*, discovered by Rayleigh,[6] who have three-color vision, but whose mixtures differ in proportion from those of normal trichromats and from one another.

Trivariance, to be expressed in the end result, must obtain at all levels of the visual process. The intrusion of a fourth variable at any point would be wasted; the reduction to two variables at any point would make the whole system dichromic. One could in principle account for dichromia through the loss of one of the three color-vision pigments; or the three pigments could be distributed between only two types of cone; or all three classes of cone could discharge their impulses through only two nerve channels; or the normal three mechanisms, distinct to that point, could arouse only two sensations. Add to these the possibility of modified color-vision pigments, and one can appreciate the wide latitude of theory that has existed in this field. Most if not all of these changes have been rung in one or another explanation of color blindness. What is more important, among all the varie-

ties of congenital and acquired color blindness, instances will probably be found in which each of these mechanisms applies. It is hardly surprising under these circumstances that some time ago it became customary to bypass theory by designating the major types of color blindness with von Kries's neutral terms *protanopia, deuteranopia,* and *tritanopia.*[7]

The key to understanding the mechanisms of color vision and its aberrations is in defining the spectral sensitivities of the color-vision pigments and the cones that contain them.[8] Stiles[9] made a fundamental contribution to this enterprise by defining with his two-color threshold technique a physiological blue-receptor mechanism that takes several forms $(\pi_1-\pi_3)$, a green-receptor mechanism, π_4, and a red-receptor mechanism, π_5. His tests showed that of these only the blue-component behaves as though it involves the action of a single photosensitive pigment.

Rushton, and, independently, Weale, undertook some years ago the very difficult problem of measuring directly the difference spectra of the red- and green-sensitive pigments through the reflectance of monochromatic lights from the fundus of the eye, having in transit passed twice through the retina ("fundus reflectometry"). Apart from the intrinsic difficulty of such measurements, their interpretation is beset with pitfalls, as Weale and Ripps[10] and Rushton[11] have lately emphasized. In a new series of papers, Rushton, working for the most part within these limitations, has offered evidence for the presence of two visual pigments in a central 2° field of the normal retina, one more red-sensitive, the other more green-sensitive.[12] A deuteranope seems to contain the former pigment alone,[11] a protanope the latter.[13] The degree to which these measurements represent the spectra of the visual pigments is left undecided.

Recently, the difference spectra of the red- and green-sensitive pigments have been measured by direct microspectrophotometry of human and monkey foveas.[14] It has proved possible also to measure these and the blue-sensitive pigment in single parafoveal cones.[15, 16] A simple psychophysical procedure has also been designed that, through selective adaptation with bright colored background lights, effectively isolates the spectral sensitivities of the single color-vision pigments.[17]

All these procedures are now in fair agreement. They show the human cones to contain three photosensitive pigments, segregated for the most part in three classes of cone, with absorption maxima (λ_{max}) near 435, 540, and 570 mμ. Owing in large part to individual differences in the yellow pigmentation of the lens and macula, the spectral sensitivities measured *in vivo* are more widely dispersed, λ_{max} appearing at 440–450, 540–550, and 565–580 mμ. The spectra of the blue- and green-sensitive pigments agree reasonably well with Stiles's π_{1-3} and π_4 mechanisms; but his π_5 clearly involves both the green- and red-sensitive pigments.[17]

There are strong *a priori* indications of the relation of these pigments to color blindness. Protanopes are abnormally insensitive to red light, as though they lack the red-sensitive pigment, and are in that sense red-blind. In the same sense, tritanopes appear to be blue-blind, and certain deuteranopes green-blind (cf. ref. 17).

It is the deuteranopes, however, who raise problems, for the spectral sensitivity of their cone vision—e.g., their photopic luminosity function— overlaps widely with the normal. Yet if they lack one of the normal color-vision pigments, that should be expected to distort their luminosity curve.

In a first attempt to understand the mechanism of color blindness on the basis of

data already in the literature,[17] I supposed that there were in reality two kinds of deuteranope, as earlier suggested by Willmer,[18, 19] one lacking the green-sensitive pigment and in that sense green-blind (a "loss" mechanism; Willmer's type II), the other possessing all three pigments in normal proportions, and hence a normal spectral sensitivity, but with matters so arranged that the green- and red-sensitive pigments excite a single sensation (a "fusion" mechanism; Willmer's type I).

With the method of selective color adaptation, one should be able to identify directly the photopigments in color-defective eyes, and to estimate their relative proportions. That is the object of the present investigation.

With the procedure standardized for this purpose, the properties of the color-vision pigments were measured in normal trichromats, all three classes of dichromat, and anomalous trichromats. In every instance, dichromats were found to lack one of the three normal color-vision pigments; no example could be found involving a "fusion" mechanism. In this procedure, anomalous trichromats yielded the same results as the corresponding classes of dichromat; here a third mechanism must be present, but much reduced in sensitivity, and displaced in spectrum. These measurements, together with what is known of the biochemistry of the visual pigments, yield some insight into the molecular and genetic mechanisms that underlie color-defective vision.

Procedures.—Color-defective subjects: Most subjects were college students, 19–25 years old. In a first screening, several hundred students were tested with the Ishihara charts[20]—mainly plates 22–27—to which we had added a tritan plate designed by Farnsworth.[21] These tests separated out the color defectives and made a first division into protans, deutans, and tritans—Farnsworth's terms for the three types of dichromat and related anomalous trichromate (i.e., protans = protanopes + protanomals; deutans = deuteranopes + deuteranomals, etc.). All the color defectives were now given the Farnsworth 15-Hue Dichotomous Test,[22] which tends to single out the true dichromats.

A number of persons who test as protans or deutans with the Ishihara plates perform normally on the Farnsworth Dichotomous Test. We suspected that these were our anomalous trichromats, and this was verified in almost every case by measurements with the Hecht-Shlaer anomaloscope.[23] In this instrument, red and green lights can be mixed in various proportions to match a yellow. Normal subjects match the yellow with a specific mixture of red and green. Protanopes and deuteranopes accept the normal mixture, but can also match yellow with a wide range of red-green mixtures including pure red or green. Anomalous trichromats reject the normal mixture, requiring either more red than normal (protanomalous) or more green than normal (deuteranomalous) to match the yellow.

Though we had no trouble finding protans and deutans by these methods, tritanopes were another matter, since they occur very rarely, perhaps one in 20,000 persons.[24] I am indebted to the Medical Research Laboratory at the U.S. Submarine Base at New London, Connecticut, Farnsworth's old laboratory, where in the course of testing many thousands of persons, mainly during World War II, they had found several tritanopes. Fortunately two of them could still be traced in the Boston area, one of whom (R. B.) had been a subject in our laboratory about 10 years before.[24] This subject, as I worked with him, told me that his 79-year-old mother has the same trouble, and she became my third tritanope.

Method of measurement: The basic procedure has already been described.[17] Visual thresholds were measured wavelength by wavelength in the dark-adapted fovea, to obtain the over-all spectral sensitivity function of the foveal cones. Then, with the eye adapted to continuous, bright-colored backgrounds, on which the test flashes were superimposed, thresholds were again measured across the spectrum. The colored backgrounds so depress the sensitivities of two of the three color-mechanisms that the third mechanism can be measured virtually in isolation. So, for example, exposing the eye continuously to a bright yellow light so depresses the sensitivity of the red- and green-sensitive mechanisms that the measurements involve the blue-receptor almost alone. Similarly, adaptation to bright blue light isolates the red-sensitive mechanism, and adaptation to wave bands in the blue and red, hence purple light, isolates the green-sensitive mechanism.

For this investigation the procedure was standardized so as to yield comparable results throughout. The apparatus was as earlier described, except that the zirconium arc formerly used as test source was replaced for greater stability with a tungsten-filament lamp.

Test conditions were as follows: 1° test field, centrally fixated and exposed for $1/25$ sec; colored background fields 3.5° in diameter and similarly centered; both test and background fields were seen in Maxwellian view, through an artificial pupil 3.5 mm in diameter.

The colored backgrounds were (1) yellow field for isolating the blue-sensitive mechanism: white field brightness 4700 ft-lamberts and color temperature 2100°K, before passing through Corning filter *3482* plus Jena heat filter *KG 1;* (2) purple field to isolate the green-sensitive mechanism: white field brightness 19,200 ft-lamberts and color temperature 2400°K before passing through Wratten filter *35;* (3) blue field to isolate the red-sensitive mechanism: white field brightness 14,900 ft-lamberts and color temperature 2300°K before passing through Wratten filter *47* plus Jena *BG 18.* The measurements are expressed as log relative sensitivity (log 1/threshold), in terms of the relative numbers of photons per flash incident on the cornea of the eye.

Observations.—Normal observers: In the standardized procedure, the intensities of the colored backgrounds were set arbitrarily so as to make the maximum sensitivities of the color-vision mechanisms about equal in the normal eye. This was done only for convenience. The correct relationships of these mechanisms in making up the average foveal luminosity function are shown in Figure 1. The red- and green-recptor mechanisms contribute almost equally whereas the blue makes only a very small contribution.[17]

Nine normal trichromats were examined with the standard procedure. The observations on three of them are shown in Figure 2, and all the data are summarized in Table 1. The data in Figure 2 exemplify the variability encountered in such normal subjects. B. K. displays close to the average relationships set by the standard procedure, C. S. is somewhat red-poor, and C. O. strikingly blue-rich, enough so as to distort the shape of his foveal sensitivity function, which is unusually high at short wavelengths. These are not the extreme variations of this kind already encountered. They extend a theme stressed earlier, that normal trichromats display large individual differences in the contributions of the three color mechanisms to the

total luminosity function. As shown earlier also, the long-wavelength shoulder that accompanies the main peak of the blue-receptor is not a property of that receptor, but residues of response from the green- and red-receptors.[17] In these normal subjects the average wavelength of maximum sensitivity (λ_{max}) of the dark-adapted fovea lies at 559 mμ, and of B, G, and R at 442, 546, and 571 mμ.

Dichromats and anomalous trichromats: Similar measurements on two protanopes and one protanomalous subject are shown in Figure 3. In all, eight protans were texmined (five protanopes and three protanomals). Both groups yielded essenialaly the same results, which are summarized in Table 1. In every instance, the procedures for isolating the blue- and green-receptors yielded normal results; but every attempt to measure the red-receptor yielded instead only the green-receptor mechanism again, at a lower level of sensitivity.

That is, only two of the three color mechanisms could be found in these subjects, whether protanopes or protanomals. The absence or near absence of the red-receptor mechanism is evident also in the shape and position of the foveal sensitivity function (D), which is abnormally narrow, owing to the falling off of sensitivity in the red, and for the same reason is displaced an average of 11 mμ toward shorter wavelengths, λ_{max} lying at 548 mμ (Table 1).

Figure 4 shows similar measurements on two deuteranopes and one deuteranomal. In all such subjects, the procedures for isolating the blue- and red-receptor mechanisms yield normal results; but every attempt to measure the green-receptor mechanism yields instead only the red-receptor mechanism at a lower level of sensitivity.

As already said, I had expected originally to find two classes of deuteranope, one lacking the green-receptor mechanism as here, the other possessing all three color-vision pigments in normal proportions, though with the red- and green-receptor mechanisms so arranged as to excite a single sensation. Not finding this second, fusion type of deuteranope among the first subjects tested, I went on to measure in all 20 deutans—13 deuteranopes and 7 deuteranomals. All of them yielded essentially the same type of result (Table 1). In all these subjects, the green-receptor mechanism is absent, or so low in sensitivity that it cannot be found by this procedure.

FIG. 1.—Contributions of the blue-, green-, and red-receptor mechanisms (B, G, R) to the over-all spectral sensitivity of the fovea. The B, G, and R curves are taken from measurements on subject R. H.[17] The over-all sensitivity is a composite of classic measurements of the photopic luminosity function (cf. ref. 17).

Measurements on the three tritanopes are shown in Figure 5. They are also summarized in Table 1, but with so few and dissimilar subjects, the averages do not mean much. In all three subjects every attempt to measure the blue-receptor mechanism failed, yielding instead curves close to the foveal luminosity function (D) at lower levels of sensitivity. The failure to find a blue-receptor is especially impressive, because this lies so far from G and R that selective adaptation isolates it very easily. Using a less rigorous procedure, we had been unable to identify a blue-receptor in the parafovea of R. B. about 10 yr earlier.[25]
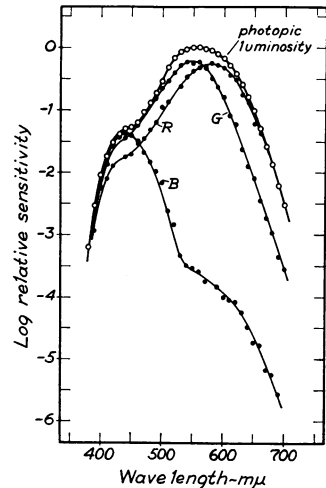
Both R. B. and his mother display relatively normal green- and red-receptor mechanisms, though the mother possesses a considerably lower visual sensitivity throughout, and is particularly insensitive in the blue and violet, presumably owing largely to the yellowing of the lens and other ocular tissues that goes with advanced age.

Subject G. G., on the other hand, yielded an apparently normal red-receptor mechanism, but the measurements that should have represented his green-receptor mechanism were almost the same, as though he possesses only the red-receptor. In that case he should have behaved as a cone-monochromat, but in fact he discrimi-
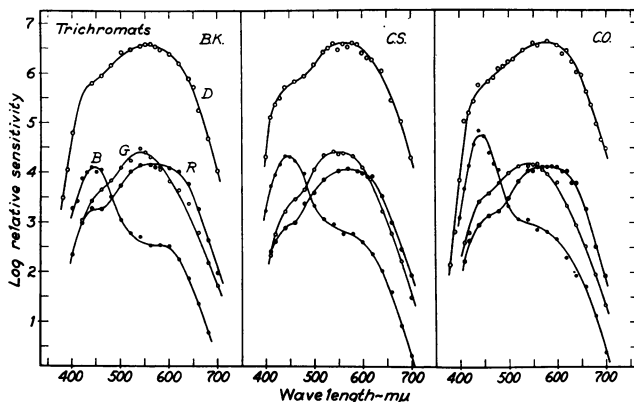

FIG. 2.—Measurements on three normal subjects, showing the spectral sensitivity of the dark-adapted fovea (D), and of the blue-, green-, and red-receptors (B, G, R) as isolated by adaptation of the fovea to intense colored backgrounds. The procedure in these experiments was standardized arbitrarily so as to make the maxima of B, G, and R about equal in the average observer. Ordinates are log relative sensitivity (log 1/threshold) expressed in terms of the relative numbers of photons per flash incident on the cornea of the eye. These data are typical, and show the variability in the proportions of B, G, and R encountered among normal subjects. B. K. is about average, C. S. a little red-poor, and C. O. markedly blue-rich, sufficiently so to distort the shape of his foveal luminosity function, which is unusually high in the blue. In nine such normal observers, the wavelength of maximum sensitivity (λ_{max}) of D lies at 559 mμ, and of B, G, and R at 442, 546, and 571 mμ.

nates colors very well. A posthumous paper by Gordon Walls[26] reported the surprising observation that three tritanopes to whom he had presented wavelengths of the spectrum in random order were able to name the colors almost perfectly. This subject can do the same.

My measurements showed this subject to lack the blue-receptor, but surely he has two others. Since I could not differentiate them by the standard procedure, I tried using intense blue and red backgrounds, with the result shown in Figure 6. A

TABLE 1

AVERAGES OF MEASUREMENTS ON NORMAL AND COLOR-DEFECTIVE SUBJECTS

Type	Number	Dark-Adapted Fovea			Blue-Receptor		Green-Receptor		Red-Receptor	
		λ_{max}	log V_{max}	log $\dfrac{V_{max}}{V_{650}}$	λ_{max}	log V_{max}	λ_{max}	log V_{max}	λ_{max}	log V_{max}
Normals	9	559	6.50	0.90	442	4.15	546	4.21	571	4.04
Protans	8	548	6.51	1.70	442	4.14	545	4.67	—	—
Deutans	20	566	6.44	0.85	442	4.13	—	—	568	4.29
Tritans	3	569	6.41	0.82	—	—	550	4.02	571	4.27

λ_{max} = wavelength of maximum sensitivity, in mμ, measured in terms of light incident on the cornea of the eye; log V_{max} = logarithm of the maximum sensitivity (1/threshold) in arbitrary units. V_{650} = sensitivity at 650 mμ in the same arbitrary units.

differentiation of color mechanisms is now evident, yet surprisingly small. As it turned out, a blue-adaptation seven times as intense as the standard did not change the shape of the curve for the red-receptor; and this, the broader curve in Figure 6, agrees very well with the average R curves of normal subjects and deutans. Adaptation to intense red light, however, yielded the narrower function shown with open circles; and this apparently represents this subject's green-sensitive pigment. It is considerably broader than the average G curve found in normal subjects and protans, and is displaced about 10 mμ toward the red, its λ_{max} lying at about 555 mμ. Also this curve is poised at a considerably lower level of sensitivity than that of the red-receptor; otherwise I should have found it with the standard purple adaptation.

Incidentally, in using intense colored backgrounds with this subject I had re-newed evidence of his excellent color discrimination. After 1 min of exposure to the intense blue field he exclaimed spontaneously that it had turned red; conversely,
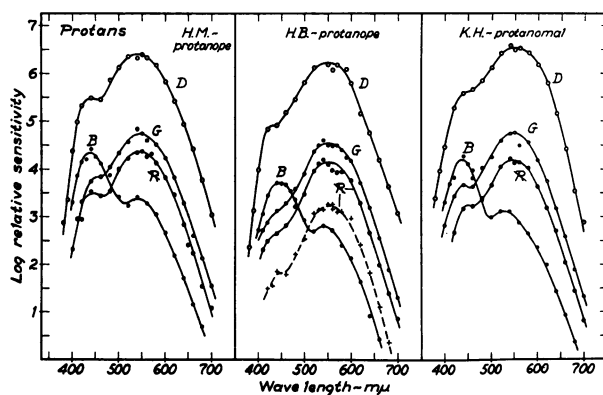

FIG. 3.—Measurements on two protanopes and one protanomalous subject. In all such subjects the standard procedure yields normal results for the blue- and green-receptors (B and G), but every attempt to isolate the red-receptor yields only the green-receptor again at lower levels of sensitivity (R with line drawn through it). With subject H. B., in addition to the standard blue-adaptation, a much higher level of blue-adaptation was tried in order to search more deeply for the red-receptor, but with no greater success (*crosses, broken line*). The lack of the red-receptor is evident in the shape and position of the spectral sensitivity curve of the dark-adapted fovea (D), which is abnormally narrow owing to the sharp falling-off of sensitivity in the red, and for the same reason is displaced markedly toward shorter wavelengths.

the intense red background turned deep green—both responses characteristic of normal observers.[27] It proved to be exceedingly fortunate to have come upon this subject, for, as will appear shortly, he makes a decisive contribution in an unex-pected direction—the mechanism of color anomaly.

To summarize the position to this point: In all these instances of color defect, such a simple loss mechanism obtains as was proposed by Thomas Young to explain Dalton's red-blindness. What is lost in each case is one of the three normal color-vision pigments. Conversely, the pigments retained in each case—except for the minor pigments associated with color anomaly, of which more below—appear to be altogether normal. The color receptors that can be measured with the standard procedure in all the dichromats and anomalous trichromats examined have the same spectral sensitivity curves as those of normal observers. Any doubt that this procedure isolates to a high degree the spectral sensitivities of the single color-vision

pigments, even when they overlap greatly as in the green- and red-receptors, should be largely allayed by this evidence that the green-receptor displays the same spectral sensitivity in the normal observer as in protanopes, where there is no red-receptor to interfere with it; and similarly the red-receptor displays the same properties in normal subjects as in deuteranopes, where no green-receptor interferes. Again the blue-receptor exhibits the same behavior in all the subjects who possess it. As stated earlier, the shoulder at longer wavelengths that accompanies the main peak of the blue-receptor is not part of this function, but represents residues of sensitivity from the green- and red-receptors.[17] For this reason, this shoulder appears at about 550 mμ, near the peak of the green-receptor, in protans (Fig. 3), but at about 570–580 mμ, near the peak of the red-receptor, in deutans (Fig. 4).

It seems from these measurements that only three major kinds of dichromat exist, depending upon which pigment is lacking. In that case, however, it would be well to drop the arbitrary terms protanope, deuteranope, and tritanope, meaningless to
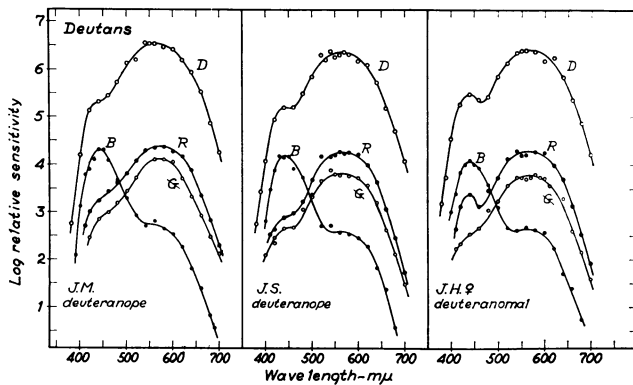

FIG. 4.—Measurements on two male deuteranopes and one female deuteranomal. In all such subjects tested, the standard procedure reveals normal blue- and red-receptors (B and R); but all attempts to measure the green-receptor yield instead only the red-receptor again at a lower level of sensitivity (G with line through it).

all outside the field, and to call the three kinds of dichromat blue-, green-, and red-blind; or if technical language is preferred, acyanopes, achloropes, and anerythropes. (The corresponding anomalous trichromats could be called blue-, green-, and red-anomals, or cyanomals, chloranomals, and erythanomals.) Having begun this paper with the arbitrary terms, I shall go on with them; but I hope it may become common practice in the future to use the more descriptive terms, more readily understood by all.

Mechanism of anomalous trichromia: Throughout these experiments I have obtained essentially the same results from anomalous trichromats as from dichromats, in each instance finding two of the three normal color receptors. Anomalous trichromats obviously possess a third receptor. What are its properties?

Two statements can be made of it: (1) It is represented poorly, in the sense of having a low sensitivity, contributing little therefore to the total luminosity function; otherwise I should have found it. (2) This in itself is not enough to account for the abnormal color matching that is characteristic of anomalous trichromats, as in the anomaloscope (Table 2). That demands, as König and Dieterici empha-

sized,[28] that at least one of the color mechanisms have a displaced spectral sensitivity, not derivable from any of the normal functions by homogeneous linear transformation (cf. also ref. 28a). This distinction must attach to the missing component in our measurements, but how do we examine it?

That consideration makes particularly important the result obtained with tritanope G. G., shown in Figure 6, for in addition to lacking the blue-receptor, his green-receptor is displaced markedly toward the red.

The usual anomaloscope, in which yellow is matched with mixtures of red and green, takes no account of the blue-receptor, and hence of tritanopia. In the anomaloscope, tritanope R. B., who has normal green- and red-receptors, yields the normal color match; but G. G., as was to be expected from the curves of Figure 6, requires far more green in the red-green mixture to match yellow (Table 2).

TABLE 2

ANOMALOSCOPE MEASUREMENTS IN NORMAL AND COLOR-DEFECTIVE SUBJECTS

Observers (nos.)	Average reading	Range
Normal (20)	47.1	44.1–49.3
Protanomalous (4)	31.7	26.0–38.7
Deuteranomalous (14)	57.4	52.4–62.2
Tritanope R. B.	47.6	47–49
Tritanope G. G.	58.5	58–59

In the Hecht-Shlaer anomaloscope, red and green lights are mixed to match a yellow. Smaller numbers than the normal matching range involve adding more red, larger numbers adding more green.

He is not only a tritanope, but anomalous, our first example of an anomalous dichromat, a deuteranomalous tritanope. As such, he seems to be showing us for the first time the mechanism of deuteranomaly, in possessing a normal red-receptor and an abnormal green-receptor, low in sensitivity, and with its spectrum displaced toward the red.
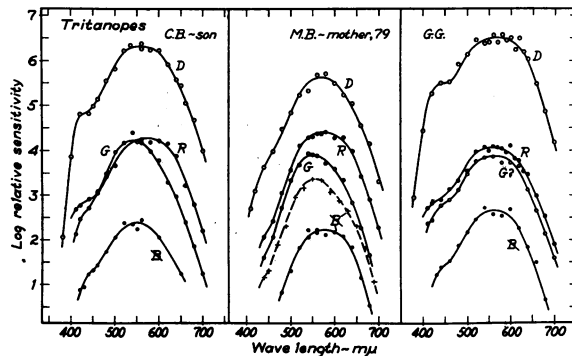

FIG. 5.—Measurements on three tritanopes. The first two subjects, son and mother, yielded elatively normal functions for the red- and green-receptors (*R, G*); but all attempts to measure the blue-receptor yielded instead approximations to the other curves at lower levels of sensitivity. All the data of M. B. are depressed in sensitivity and narrowed by a special falling-off at shorter wavelengths, probably owing to the deepening of the yellow pigmentations of the lens and macula that go with advanced age. With M. B., in addition to the standard yellow adaptation, the same light was tried at considerably lower intensity, in the hope of revealing some blue-response (*crosses, broken line*), but without success. The third subject, G. G., also yielded no blue-response; but here whereas the standard procedure for isolating the red-receptor yielded a relatively normal result, the attempt to isolate the green-receptor yielded essentially the same result, as though this subject possessed only the red-receptor. He had, however, excellent color discrimination, and proved to be deuteranomalous as well as a tritanope (see Fig. 6).

A possible biochemical basis for the displaced spectral sensitivity curves of color-anomalous observers is discussed below. However, we can dispose here of one possibility that has been frequently considered: that differences in the density of macular pigmentation could have such effects. This can be tested directly, for the human macular pigment is lutein or leaf xanthophyll, $C_{40}H_{54}(OH)_2$;[29] and we have made up for use as a color filter a solution of this pigment in a mixed solvent (2% aqueous digitonin and methanol), the absorption spectrum of which mimics closely that of the macular pigmentation measured directly in human retinas.[30, 17] When this is placed behind the eyepiece of the anomaloscope, its effect is added to the normal macular pigmentation, the average absorbance of which is about 0.5 at the maximum, 460 mμ. When the added filter has an absorbance 1.0—i.e., 90 per cent absorption at 460 mμ—it has no appreciable effect on anomaloscope readings. One can calculate that for the macular pigmentation to cause the abnormal anomaloscope readings encountered in deuteranomalous observers would require macular absorbances at 460 mμ of at least 10.0, a wholly absurd possibility. Furthermore, protanomaly, which apparently involves some displacement of the red-sensitive pigment toward shorter wavelengths, would in any case demand another kind of explanation (cf. also ref. 28a).
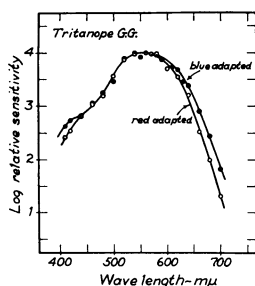

FIG. 6.—Spectral sensitivity measurements on tritanope G. G., adapted to intense blue and red lights. The blue light was that used in the standard blue adaptations, but about seven times as bright. The red adaptation involved a white source of brightness 166,000 ft-lamberts and color temperature 2920°K passing through Jena filters *RG 5* and *KG 1*. This extreme treatment reveals only small differences in spectral sensitivity. The wider curve (blue-adapted) is the normal red-receptor. The narrower curve (red-adapted) represents this subject's green-receptor. Relative to the normal, it is displaced toward the red and is low in sensitivity, making this subject deuteranomalous.

Discussion.—Foveal luminosity functions: Figure 7 compares certain features of the spectral sensitivity curves of the dark-adapted fovea in the various types of observer, and pursues further the question whether there are two classes of deuteranope.

In the upper part of Figure 7, λ_{max} is plotted for all the classes of subjects. Its distribution in normals overlaps only slightly with that in protans, but very extensively with that in deutans. Yet the deutan distribution is very different from normal, the average λ_{max} in these measurements lying 7 mμ further toward the red (cf. refs. 31 and 32). This is only slightly smaller than the departure of the protan average from normal (11 mμ; Table 1).

The lower part of Figure 7 shows the distributions of an index for the *shape* of the luminosity function—the logarithm of the ratio of the sensitivity at the maximum to that at 650 mμ. This index has about the same distribution in normals, deutans, and tritanopes, but is very different in protans, in whom the luminosity curve is markedly narrowed by the sharp decline of sensitivity in the red.

It is striking that with regard to both λ_{max} and shape, the luminosity functions of protanomalous subjects are hardly to be told apart from those of protanopes, and

those of deuteranomals hardly to be distinguished from those of deuteranopes (cf. refs. 32 and 33).

Neither type of plot in Figure 7 shows any tendency on the part of the deutan data to be bimodal. The measurements on deutans spread widely, whatever pa-
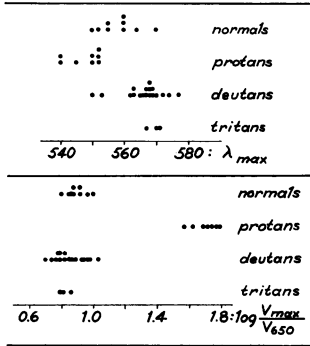

FIG. 7.—Distribution of indices for the position and shape of the spectral sensitivity function of the dark-adapted fovea in normal and color-defective subjects. *Upper:* distribution of the wavelength of maximum sensitivity (λ_{max}). In protans this overlaps only slightly with the normal distribution, but in deutans, though the average λ_{max} lies at distinctly longer wavelengths, there is a large overlap with the normal distribution. *Lower:* the logarithm of the ratio of the sensitivity at the maximum to that at 650 mμ is taken as an index of the shape of the spectral sensitivity function. By this criterion the shapes of the normal, deutan, and tritan functions are very similar; but the protan function is abnormally narrow, owing to the sharp decline of sensitivity in the red. Neither of these indices shows any tendency toward bimodality in deutans.

rameters are considered, but they remain essentially homogeneous. There is no encouragement in any of these measurements for dividing deutans—or among them the deuteranopes—into two groups.

That raises again the question how the deutan luminosity curve—granted its average displacement toward the red—resembles so greatly in shape and position the normal function. Why is it not distorted as is the protan luminosity curve? I think the answer is disarmingly simple. It is that though the green-receptor, which accounts principally for the protan luminosity curve, fails badly in the red, the red-receptor, which mainly accounts for the deutan luminosity curve, does not fail correspondingly in the green. On the contrary, it makes a surprisingly close approximation to the over-all spectral sensitivity of the normal fovea.

This realization has a curious history. König, reconsidering in 1897[8] the results of his monumental investigation with Dieterici of normal and defective color vision,[28] proposed a revised trio of "basic sensation curves" (*Grundempfindungen*) which, when corrected for the energy distribution of his source, agree astonishinngly well with present measurements.[34] In retrospect one sees also that the red-component of this trio came close together with König's measurements of the photopic luminosity function. König seems never to have made this comparison in his writings; but long after his death two of his former students, von Kries[35] and Ladd-Franklin (cited by Judd[34]), alleged that he was not only well aware of it, but had concluded for this reason that the red-component accounts nearly or entirely for the luminosity of the spectrum, the violet- and green-receptors contributing to hue but hardly, if at all, to luminosity. This view is now at times ascribed directly to König, I think with some injustice. Obviously it would account easily for those deuteranope luminosity curves that overlap with the normal; but what of the others that are displaced toward the red? And how are protans to have a luminosity function if the green-receptor does not provide it? Yet a point remains: the spectral sensitivity of the red-receptor alone has nearly the same shape and overlaps widely in position with the total cone luminosity function.

Have dichromats a lowered absolute sensitivity? Some years ago Hecht and Hsia[36] reported that certain dichromats whom they had examined exhibited absolute losses of sensitivity throughout the spectrum compared with normal observers. The maximum sensitivity of the dark-adapted fovea was reported to be on the average 0.21 log unit lower than normal in seven deuteranopes, and 0.24 log unit lower than normal in six protanopes. Normal individuals vary so widely in this regard, however, that much larger numbers of dichromats would have been needed to establish the reality of such small differences.

The present measurements contain no clear indication of lowered maximum sensitivity in the dark-adapted foveas of dichromats or anomalous trichromats as compared with normals (Table 1). Protans display absolute losses of sensitivity in the red, and many deutans are noticeably low in sensitivity in the green and blue, as reported earlier by Hsia and Graham;[31] but those are direct and obvious consequences of the distortions and displacements of the luminosity function in such observers already considered.

Fate of the missing cones: A closely related consideration involves the fate of the missing class of cones in dichromats. It would be of the greatest value to have histological studies of dichromic eyes, with counts of rods and cones and their central connections; but I do not know of any such material. There is no present reason to believe, however, that dichromats have fewer foveal cones than normal; and their apparently normal visual sensitivity and acuity are evidence to the contrary.

If dichromats do possess normal numbers of cones, the missing category must be replaced by others, the red-cones lacking in protanopes presumably being replaced mainly by green-cones, and the green-cones lacking in deuteranopes mainly by added numbers of red-cones. The situation should be much the same in anomalous trichromats, in which the underrepresentation of the aberrant class of cone should demand similar substitutions. That might mean approximately doubling the numbers of green-cones in protans and of red-cones in deutans. Doubling the numbers of cones should have much the same effect as doubling the area of the test field; i.e., it might approximately halve the threshold or double the sensitivity, so raising log V_{max} by about 0.3. In Table 1, the average log V_{max} of the green-cones is 0.46 higher in protans than in normals, and the log V_{max} of the red-cones is 0.25 higher in deutans than in normals. These are small differences, but of about the expected size. The comparison is made more significant by the fact that in all three classes of subject log V_{max} remains almost identical for the blue-receptor, so providing an internal reference point, relative to which the sensitivities of the green-receptors in protans and the red-receptors in deutans are increased by roughly the expected amounts. These differences appear therefore to be real, and to imply the substitutions considered above: that the red-cones lost by protans are replaced mainly by increases in the numbers of green-cones, and the green-cones lost by deutans are replaced mainly by added red-cones.

Genetics of defective color vision:[37] Every schoolboy learns that color blindness is inherited as a sex-linked recessive; it is the archetype of this kind of inheritance in man.[38] That holds only for the protan and deutan conditions, however, and even there the genetics still presents problems. So, for example, it has remained undecided whether one or two loci on the X chromosome are involved.[37, 39]

The demonstration that the human red- and green-sensitive pigments, after

bleaching in the light, can be regenerated by the addition of 11-*cis* retinaldehyde shows that this is their common chromophore, and that these pigments differ only in their protein component, their opsins.[14]   That two different proteins are involved in forming these pigments is strong presumptive evidence for the operation of two different genes.

The simplest assumption is that two loci on the X chromosome govern the synthesis of the opsins of the normal red- and green-sensitive pigments.   A mutation at either locus accounts for the loss of the corresponding pigment, hence protanopia or deuteranopia.   Protanomaly and deuteranomaly also both breed true, and seem to represent other mutations at the same two loci, recessive to the normal but dominant to the color-blind condition.[39]   That is, at each locus at least three alleles represent the normal, color-anomalous, and color-blind conditions, with genetic dominance descending in that order.   It is very likely that more detailed measurements will reveal further differences within the present categories of color anomaly; so Franceschetti has proposed that "extreme" color anomaly represents a fourth allelic condition.[40]

Tritanopia has a quite different inheritance.   It is a rare condition, the most optimistic estimate assessing its frequency at about 1 in 13,000 persons.[41]   Associated with Wright's investigation of tritanopia in England,[24] Kalmus studied its familial distribution.[41]   Of 47 tritanopes included in this study, 18 (38%) were females. This appears therefore to be not sex-linked, but an autosomal trait.

In the immediate pedigree of my own subjects, tritanopia appears suddenly and in apparent isolation.   The mother-and-son pair M. B. and R. B. seem to be the only members of their immediate families to show this trait.   R. B. is an only child; and his only child, a girl, was tested and is normal.   His father and all his mother's and father's siblings are said to have been normal.   Tritanope G. G. also appears to be the only such case in his family.   I tested his father, his son and daughter, his brother, and the brother's two children, and found all of them normal.

I thought originally that tritanopia may represent an autosomal recessive, for its incidence is much like that of protanopia or deuteranopia *in women* (i.e., where two doses of the recessive gene are needed to develop the trait).   However, Kalmus concludes that his pedigrees are more consistent with the irregular manifestation of one or more autosomal dominants.   The situation is still far from clear, not surprising in view of the small number of cases yet available.   Kalmus's pedigrees exhibit curious irregularities; so, for example, six tritanopic females had ten tritanopic sons but only one tritanopic daughter (in addition to one normal son and one normal daughter)—a strange sex ratio, if nothing else, to which can be added my own tritanopic mother and only son.   The rarity and sporadic appearance of tritanopia might be ascribed to low "penetrance," i.e., the trait, though carried genetically, frequently is not manifested.   Low penetrance, however, is usually associated with indirect, secondary effects of mutation.   So, for example, the gene for phenylketonuria exhibits low penetrance in its effect upon head size or hair color, but is manifested completely if one measures phenylpyruvate in the urine.   To ascribe low penetrance to tritanopia would suggest that this also is a remote consequence of mutation, rather than the direct failure to form the opsin of the blue-sensitive pigment.

I think it very likely that what is now called tritanopia includes several different conditions. It is perhaps one sign of this that of my three tritanopes, two forming a mother-and-son pair, the third displays the aberrant result discussed above. Kalmus remarks on another complication: it is sometimes hard to tell tritanopia from tritanomaly, which seems to be inherited as a sex-linked recessive.[42, 41] In that case the genes for normal blue vision, tritanopia, and tritanomaly do not form such an allelic series as do those concerned with the red- and green-receptor mechanisms and their anomalies.

Blue-blindness and monochromia: Among the properties of the blue-receptor mechanism, two stand out: its poor representation in normal vision (cf. Fig. 1), and the low incidence of congenital blue-blindness. These seem to lead directly to two consequences: (1) Any *functional* attack on cone vision is likely to result in the loss of the blue-receptor, while leaving adequate representations of the red- and green-receptors. This may be why degenerative diseases of the retina and retinal detachment frequently lead to acquired tritanopia.[8, 43, 44] (2) Conversely, *genetic mutation* is more likely to affect the red- and green-receptor mechanisms than the blue. That may be why numbers of so-called "typical" monochromats— monochromats possessing poor visual acuity—retain the blue-receptor mechanism alone.[45, 46] (This is not the whole story, for other "typical" monochromats may have only rods; and still others have rods and cones, both of which display rodlike spectral sensitivities.[47, 48]) It seems probable that in those monochromats who possess only the blue-receptor mechanism, two mutations have resulted in the simultaneous loss of the more mutable red- and green-receptor mechanisms, leaving only the blue. This view is borne out by the observation that this type of monochromia is inherited as a sex-linked recessive, as are the protan and deutan conditions.[46, 49] The poor representation of the blue-receptor may then account for the low visual acuity of such monochromats. Indeed, a simple-minded view of monochromia would suggest that this being one mechanism of the "typical" form, the rarer, "atypical" or cone-monochromia[50, 51] involves the simultaneous loss of the blue-receptor mechanism and one other, so that only the red- or the green-receptor mechanism is retained, with its characteristically high sensitivity and visual acuity.

Molecular basis of defective color vision: It seems useful now to bring the genetics of defective color vision together with its biochemistry and physiology to make a simple hypothesis. The usual business of a gene is to specify the amino acid sequence of a protein. Presumably the genes concerned with normal color vision determine in this way the amino acid sequences of three opsins. The usual effect of mutation is to substitute one amino acid for another in such a sequence. Probably this is the effect of the mutations for color defects on the opsins. In some cases such a substitution blocks the synthesis of a visual pigment; the result is then color blindness. In other cases different amino acid substitutions result in the synthesis of modified visual pigments in abnormally small amounts; the result is then color anomaly. I see no reason to doubt that a variety of amino acid substitutions could lead to either type of result; the associated mutations would then constitute multiple alleles for color blindness and for various degrees and types of color anomaly. Both kinds of defect, if this is their genetic mechanism, should leave the other visual pigments unaffected, as we find.

Summary.—The spectral senstivities of the color-vision pigments can be isolated

and their proportions estimated by selective adaptation of the eye to intense colored lights.   In normal subjects this procedure reveals the operation of three pigments, the proportions of which vary greatly in different individuals.   In each of the three major classes of color-blind persons (dichromats), one of these pigments is lacking, though the other two remain normal.   Anomalous trichromats yield the same results in this procedure as dichromats;   here a third pigment is present, though reduced in amount and displaced in spectrum.   In one instance—that of a deuteranomalous tritanope—this mechanism could be analyzed directly.   Genes at two loci on the X chromosome specify the two opsins that join with the common chromophore, 11-*cis* retinaldehyde, to yield the normal red- and green-sensitive pigments.   Mutations in either of these genes may block the synthesis of a pigment, resulting in color blindness, or may lead to the synthesis of reduced amounts of a modified pigment, resulting in color anomaly.

Note added in proof: The suggestion that the red-cones lost by protans are replaced mainly by increases in the numbers of green-cones, and the green-cones lost by deutans are replaced mainly by added red-cones, may well be correct, but the argument is indecisive: (1) Doubling the concentration of one type of foveal cone may approximately double the sensitivity;   but there is no analogy with doubling the area of a centrally fixated foveal test field, which raises the sensitivity more nearly 1.3–1.4 times, in line with the fact that the density of cones falls off sharply from the center of the fovea toward its rim [Østerberg, G., *Acta Ophthalmol. Suppl.* 6 (1935)], so that doubling the area does far less than double the number of cones in the field.   (2) If the missing cones are replaced by others, why not also by blue-cones; i.e., why doesn't the sensitivity of the blue-mechanism increase proportionately with that of the red-mechanism in deutans or the green-mechanism in protans, rather than remaining constant?   (3) There is room for an added or alternative consideration: that in the normal eye the red- and green-mechanisms mutually inhibit each other, as already demonstrated electrophysiologically in a number of situations, though not yet for a primate fovea.   In that case in a protan, relief from inhibition by the red-mechanism should raise the sensitivity of the green-mechanism;   and conversely in a deutan, relief from inhibition by the green-mechanism should increase the sensitivity of the red-mechanism.

¹ Dalton, J., *Mem. Lit. Phil. Soc. Manchester*, **5**, pt. I, 28 (1798).

² Young, T., *Lectures on Natural Philosophy and the Mechanical Arts* (London: Joseph Johnson, 1807), vol. 2, p. 345.

³ Wilson, G., *Researches on Colour-Blindness* (Edinburgh: Sutherland-Knox, 1855), p. 106.

⁴ Maxwell, J. C., in Wilson, G., *Researches on Colour-Blindness* (Edinburgh: Sutherland-Knox, 1855), p. 153.

⁵ Maxwell, J. C., *Scientific Papers* (Cambridge: Cambridge University Press, 1890), vol. 1, p. 121.

⁶ Rayleigh, J. W. S., *Nature*, **25**, 64 (1881).

⁷ von Kries, J., in von Helmholtz, H., *Treatise on Physiological Optics*, ed. J. P. C. Southall (Washington, D.C.: Optical Society of America, 1924), vol. 2, p. 402.

[8] König, A., *Sitzber. Akad. Wiss. Berlin* (1897), pp. 718–731: "The final goal of the theoretical investigation of color consists *inter alia* in this: the determination of the spectral sensitivities of the basic sensations."

[9] Stiles, W. S., *Ned. Tijdschr. Natuurk.*, **15**, 125 (1949); Stiles, W. S., these Proceedings, **45** 100 (1959).

[10] Weale, R. A., *Doc. Ophthalmol.*, **19**, 252 (1965); Ripps, H., and R. A. Weale, *Nature*, **205**, 52 (1965).

[11] Rushton, W. A. H., *J. Physiol.*, **176**, 24 (1965).

[12] Rushton, W. A. H., *Nature*, **206**, 1087 (1965); Baker, H. D., and W. A. H. Rushton, *J. Physiol.*, **176**, 56 (1965).

[13] Rushton, W. A. H., *J. Physiol.*, **168**, 345 (1963).

[14] Brown, P. K., and G. Wald, *Nature*, **200**, 37 (1963).

[15] Marks, W. B., W. H. Dobelle, and E. F. MacNichol, Jr., *Science*, **143**, 1181 (1964).

[16] Brown, P. K., and G. Wald, *Science*, **144**, 45 (1964).

[17] Wald, G., *Science*, **145**, 1007 (1964).

[18] Willmer, E. N., *J. Physiol.*, **110**, 422 (1949).

[19] Willmer, E. N., *Doc. Ophthalmol.*, **9**, 252 (1955).

[20] Ishihara, S., *Tests for Colour-Blindness* (Tokyo: Kanehara Shuppan Co.; London: H. K. Lewis and Co., 1960), 15th ed.

[21] Farnsworth, D., "Tritan test plate," supplied by Medical Research Laboratory, U.S. Naval Submarine Base, New London, Conn.; described as Chart F.2 by Kalmus and reproduced as frontispiece (ref. 37d).

[22] Farnsworth, D., *Farnsworth Dichotomous Tests for Color Blindness*, Panel D-15 (New York: Psychological Corp., 1947).

[23] Willis, M. P., and D. Farnsworth, *Reports Med. Res. Lab.*, U.S. Naval Submarine Base, New London, Conn., vol. 11, Report No. 190 (1952).

[24] Wright, W. D., *J. Opt. Soc. Am.*, **42**, 509 (1952).

[25] Auerbach, E., in *Mechanisms of Colour Discrimination* (London: Pergamon, 1960), p. 177.

[26] Walls, G., *Vision Res.*, **4**, 3 (1964).

[27] Auerbach, E., and G. Wald, *Am. J. Ophthalmol.*, **39**, No. 2, Pt. II, 24 (1955).

[28] König, A., and C. Dieterici, *Z. Psych. Physiol. Sinnesorgane*, **4**, 241 (1892). Also in König, A., *Physiologischen Optik* (Leipzig: Barth, 1903), p. 214.

[28a] Pitt, F. H. G., *Med. Res. Council, Spec. Rept. Ser.*, **200**, 51–55 (1935): ". . . anomalous trichromatism is not a uniformly reduced form of normal trichromatism. If it were caused by a uniform reduction of one or more of the fundamental sensation curves, a match for the normal trichromat must of necessity be a match for the anomalous trichromat.

"Excessive variation of macular pigmentation . . . cannot possibly explain anomalous trichromatism . . . ."

[29] Wald, G., *Science*, **101**, 653 (1945).

[30] Brown, P. K., and G. Wald, *Nature*, **200**, 27 (1963).

[31] Hsia, Y., and C. H. Graham, these Proceedings, **43**, 1011 (1957).

[32] Grützner, P., *Arch. Ophthalmol.*, **164**, 411 (1962).

[33] McKeon, W. M., and W. D. Wright, *Proc. Phys. Soc. London*, **52**, 464 (1940).

[34] Judd, D. B., *J. Opt. Soc. Am.*, **35**, 196 (1945).

[35] von Kries, J., footnote on p. 412 to a note specially prepared for the English edition of von Helmholtz, H., *Physiological Optics* (Washington, D.C.: Optical Society of America, 1924), vol. 2.

[36] Hecht, S., and Y. Hsia, *J. Gen. Physiol.*, **31**, 141 (1947–1948).

[37] The present status of the genetics of defective color vision and references to the literature will be found in: (a) Stern, C., *Human Genetics* (San Francisco: Freeman, 1960), 2nd ed.; (b) François, J., *Heredity in Ophthalmology* (St. Louis: C. V. Mosby, 1961), pp. 410–434; (c) Franceschetti, A., J. François, and J. Babel, *Les Hérédo-Dégénérescences Chorio-Rétiniennes* (Paris: Masson, 1963), vol. 2, pp. 1303–1391; (d) Kalmus, H., *Diagnosis and Genetics of Defective Colour Vision* (Oxford: Pergamon Press, 1965).

[38] Wilson, E. B., *Arch. Mikr. Anat.*, **77**, 249 (1911).

[39] Waaler, G. H. M., *Z. Induktive Abstammungs Vererbungslehre*, **45**, 279 (1927).

[40] Franceschetti, A., *Schweiz. Med. Wochschr.*, **58**, 1273 (1928).

[41] Kalmus, H., *Ann. Human Genet.*, **20**, 39 (1955).

[42] Hartung, H., *Klin. Monatbl. Augenheilk*, **76**, 229 (1926).

[43] Sloan, L. L., and D. J. Brown, *Am. J. Ophthalmol.*, **54**, 629 (1962).

[44] Verriest, G., *J. Opt. Soc. Am.*, **53**, 185 (1963).

[45] Blackwell, H. R., and O. M. Blackwell, *Vision Res.*, **1**, 62 (1961).

[46] Alpern, M., and B. E. Spivey, *Arch. Ophthalmol.*, **74**, 334 (1965).

[47] Sloan, L. L., *J. Opt. Soc. Am.*, **44**, 117 (1954).

[48] Alpern, M., H. F. Falls, and G. B. Lee, *Am. J. Ophthalmol.*, **50**, 996 (1960).

[49] Spivey, B. E., *Arch. Ophthalmol.*, **74**, 327 (1965).

[50] Pitt, F. H. G., *Nature*, **154**, 466 (1944).

[51] Weale, R. A., *J. Physiol.*, **121**, 548 (1953).