AN OPPONENT–PROCESS THEORY OF COLOR VISION

LEO M. HURVICH [1] AND DOROTHEA JAMESON [1]

Eastman Kodak Company

The two major theoretical accounts of color vision are those classified as the Young-Helmholtz and the Hering types of theories. For many years the former has been judged by most workers in the field to provide the simplest explanation of the way in which light stimuli give rise to color sensations. The advantages that appear to favor the Young-Helmholtz three-component hypothesis are two: it is parsimonious, and its postulates are easily quantifiable and hence subject to precise experimental test. In its parsimonious and easily quantifiable form, the theory is simple: in addition to the rods which subserve twilight vision, the eye contains three kinds of cone photoreceptors; each type of cone contains a differently selective photochemical substance; each is associated with its own specific nerve fiber; and each cone-photochemical-nerve fiber system is correlated with one of the three specific "fundamental" color sensations, namely, red, green, and blue (or violet). All sensations are considered as compounded of varying amounts of these three excitatory systems, with white arising from equal and simultaneous excitation of all three, and yellow from equal red and green excitations.

The Young-Helmholtz three-cone, three-nerve, three-sensation theory derives directly from the basic fact of color mixture, namely, that all visible hues can be matched by the mixture, in proper proportions, of only three physical light stimuli. Based squarely on this fact, the theory is readily quantified in terms of the three measurable variables of color mixture experiments. But the three measured variables, it

[1] Now at New York University.

must be emphasized, are the three physical light stimuli used in the color mixture experiments; they are not the postulated three "fundamental" color sensations, for with each different stimulus triad used for color matching a different and equally valid triad of color mixture functions is obtained. Consequently, throughout some hundred years since the original formulation of the idea, a continued series of attempts has been made to find the proper transformation of the three measured color-mixture curves that will bridge the gap and yield the unique spectral distribution curves of the desired physiological correlates of the three postulated "fundamental" color sensations. An infinity of such transformations is available for trial, and almost every serious adherent of the theory has proposed at least one new set of "fundamental sensation curves" (48, pp. 368–372). The search, however, continues, because serious defects have been found in every proposal made thus far. When the explanatory or predictive power of the theory in any given quantified form is tested it cannot handle more than a limited number of facts satisfactorily (11, p. 805).

Moreover, some facts of color experience seem unassimilable into the framework of the simple Young-Helmholtz theory with its three independent, fundamental, process-sensation systems. How can this system of three independent processes be made to account, for example, for the apparent linkages that seem to occur between specific pairs of colors as either the stimulus conditions or the conditions of the human observer are varied? Why should the red and green hues in the spectrum predominate

at low stimulus levels, and the yellow and blue hue components increase concomitantly as the spectrum is increased in luminance (43)? Why, as stimulus size is greatly decreased, should discrimination between yellow and blue hues become progressively worse than that between red and green (4, 10)? Why should the hues drop out in pairs in instances of congential color defect, or when the visual system is impaired by disease (29, 31)? On the other hand, since the sensation of white is granted no special physiological process in this parsimonious theory, but occurs as the fusion product of three equally large fundamental hue sensations, how account for the large degree of independence of white and chromatic qualities when the adaptation of the visual system is varied (37, 41)?

As more and more *ad hoc* hypotheses are added to the original Young-Helmholtz formulation in order to answer these and other problems forced by the increasing accumulation of experimental data, we naturally find the formulation becoming less and less precise and quantifiable, and obviously less parsimonious. We also find, however, that exactly those phenomena that require modification and extension of the simple "three-color theory" remind us more and more of its chief theoretical rival, the Hering theory of three paired, opponent color processes.

In view of this situation, it seems highly desirable that we take a close second look at Hering's alternative approach to an understanding of color phenomena. The vast accumulation of psychophysical data for which any adequate theoretical proposal must account requires that the basic postulates of the theory, as outlined qualitatively by Hering (13, 14), be restated in quantitative terms for such a critical scrutiny to be most meaningful. This paper will review our attempt to provide such a

quantitative restatement, and will summarize briefly some of the critical comparisons between the theoretical deductions and relevant psychophysical data. (Detailed quantitative accounts are given in 21, 22, 23, 25, 26, 27.)

BASIC SCHEMA FOR THE HERING THEORY

The Three Variables

The Hering theory is like the Young-Helmholtz theory in that it, too, postulates three independent variables as the basis for color vision, but the Hering variables are three pairs of visual processes directly associated with three pairs of unique sensory qualities. The two members of each pair are opponent, both in terms of the opposite nature of the assumed physiological processes and in terms of the mutually exclusive sensory qualities. These paired and opponent visual qualities are yellow-blue, red-green, and white-black.

The basic schema for the opponent-colors mechanism is shown diagrammatically in Fig. 1. The three paired opponent response systems are labeled y-b, r-g, and w-bk. The convention of positive and negative signs is used to
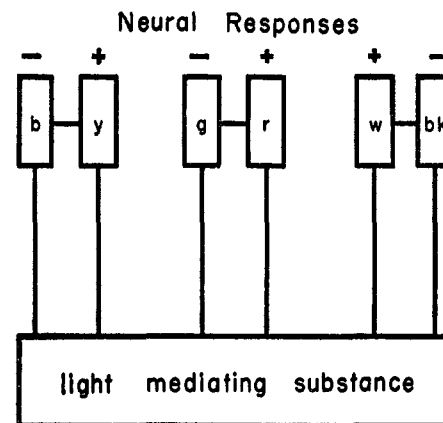

FIG. 1. Basic schema for Hering opponent-colors mechanism.

indicate that each neural system is capable of two modes of response that are physiologically opponent in nature, and that the paired sensory qualities correlated with these opposed modes of response are also mutually opponent or exclusive. That is, we may experience red-blues or green-blues but never yellow-blues, and we see yellow-greens or blue-greens, but never red-greens, and so on. In the absence of any external visual stimulus, the state of the visual system is assumed to be a condition of active equilibrium, and this equilibrium condition is associated with the neutral, homeogeneous "gray" sensation perceived after a long stay in complete darkness. This sensation is quite different from the black experience of the white-black opponent pair. Blackness arises neither by direct light stimulation nor in the simple absence of light, but rather by way of either simultaneous or successive contrast during, or following, light stimulation of some part of the retina.

Properties of Paired Systems

The three pairs of visual response processes are independent of each other; that is, they have different response thresholds, they follow different laws of increase with increase in strength of stimulation, and probably have different time constants. The achromatic system is the most sensitive; that is, the amount of photochemical absorption necessary to excite the achromatic white response is less than the amount of photochemical activity required to stimulate either the y-b or r-g chromatic pairs. This characteristic accounts for the existence of the so-called achromatic interval, i.e., the fact that spectral lines appear achromatic at the absolute threshold for visibility (42, p. 167). Similarly, the red-green system has a lower threshold than the yellow-blue one. The failure of the yellow-blue system to respond at near-

threshold levels that are sufficient to activate the red-green system exhibits itself in the facts of so-called "small field dichromasy," in which the eye behaves, with respect to stimuli that are very small in area as well as of low intensity, in a manner similar to the congenital tritanope, i.e., a specific type of "color blind" individual for whom yellow and blue discriminations are impossible and the only hues seen are reds and greens (4, 49).

With increase in level of stimulation the different paired systems also show differences in rate of response increase, such that the achromatic response increase is probably the most rapid of the three, with the result that at very high intensities all spectral stimuli show a strong whitening, or desaturation, relative to their appearance at some intermediate luminance level (42, p. 168). Of the two chromatic pairs, the yellow-blue system, although exhibiting a higher threshold, shows a more rapid rate of increase in response with increase in luminance than does the red-green system. Thus, the mixed hues of the spectrum—the violets, blue-greens, yellow-greens, and the oranges—all vary systematically with increase in spectral luminance, and all show a tendency to be more blue or yellow, respectively, at high luminances, and more red or green at the lower luminance levels (the Bezold-Brücke hue shift phenomenon).

The opponent systems show a tendency toward restoring the balanced equilibrium condition associated with the neutral "gray" sensation. Thus excitation, say, of the r process in the r-g system results in a decrease with time in r responsiveness, and in an increase in the responsiveness of the opponent g process. If we think of the r process as perhaps associated with the building up of an electrical potential in the neural system, and of the g process as associated with the collapse of the po-

tential during impulse firing, then it is easy to see that as the neural potential is increased to higher values there will be a tendency to resist further build up, and also an increased disposition of the tissue toward impulse firing in order to restore the potential to its normal equilibrium value. Although we are not at all ready to ascribe a specific neural correlate of this sort to the postulated opponent processes at this time, the neurophysiological parallels are useful for conceptualizing the opponent-process notion as a real biological phenomenon.

To return to our example, if the responsiveness of the opponent g process tends to increase as r excitation is continued, then when the r stimulus is removed we can expect g activity to be released, strongly at first, then more slowly, and ultimately fading out as equilibrium is again approached. The sensory correlate of this reversal of opponent activities with removal of stimulation is, of course, the familiar phenomenon of the complementary afterimage. If the stimulus (of constant magnitude) is not removed but continues to act for a considerable length of time, then the r process, whose responsiveness is being continuously decreased, will eventually cease to respond further, and a new equilibrium state will be reached. The disappearance of a sensory response with continued constant stimulation can be observed either by the *Ganzfeld* technique, in which the whole retina is uniformly illuminated by diffuse light (18), or by the "painted image" technique, in which optical means are used to fix a well defined image on the retina in such a way that its retinal position remains constant and independent of eye movements (39). By either method the eventual result of continued steady stimulation is a disappearance of the visual experience: the light seems to have gone out in the

Ganzfeld situation, or, in the fixed-image situation, the perceived object simply fades out of view.

Not only are the visual responses modified by changes in time in the excitabilities of the opponent processes, but they are also importantly affected by spatial interaction among the various elements of the visual field. Within certain limits there is evidence of summation of similar kinds of activity in adjacent elements, as in threshold responses for small stimulus areas (5, pp. 846–852). But perhaps more important for the over-all functioning of the visual system are the antagonistic interactions, such that r activity in one area induces g activity in adjacent areas, and similarly for the yellow-blue and white-black paired response systems. These opponent spatial induction effects are evident in all the familiar color and brightness contrast phenomena (35, pp. 138–142). They are probably also primarily responsible for the great visual-image clarity that characterizes vision in spite of the fact that the optical system of the eye is obviously imperfect, and that consequently the light image formed on the retinal surface lacks sharply defined boundaries (17, pp. 151–159). The spatial interaction causing intensification of opponent qualities at adjacent surfaces would seem an ideal crispening device to sharpen up the initially blurred retinal image.

Photochemical Postulates

In addition to the various temporal and spatial induction effects, which are assumed to be based in the neural visual-response tissue, visual adaptation probably also involves changes in the photochemical activities that initiate the neural responses, since a certain amount of photochemical bleaching is expected to occur with continued exposure of the photosensitive materials to a retinal light stimulus. In order
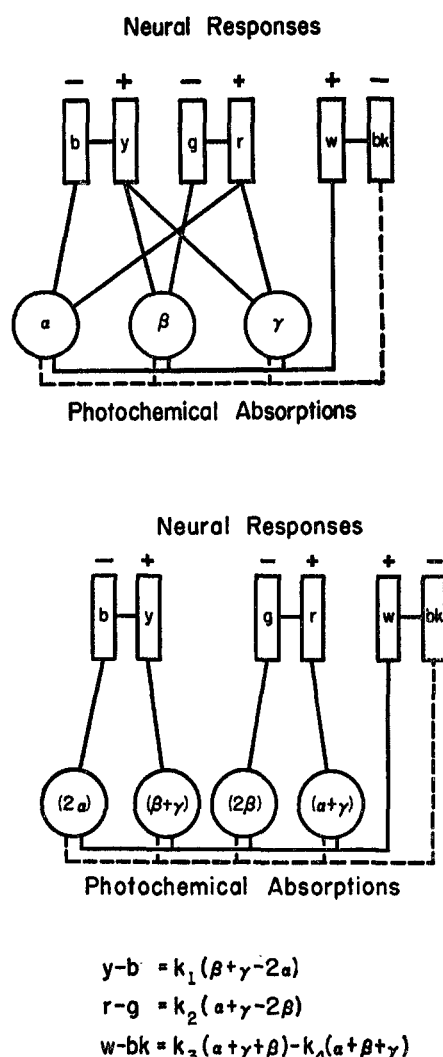
Neural Responses

Photochemical Absorptions

Neural Responses

Photochemical Absorptions

$$y-b = k_1(\beta + \gamma - 2\alpha)$$
$$r-g = k_2(\alpha + \gamma - 2\beta)$$
$$w-bk = k_3(\alpha + \gamma + \beta) - k_4(\alpha + \beta + \gamma)$$

FIG. 2. Schematic diagram showing relations between photosensitive materials α, β, and γ and neural opponent response processes y-b, r-g, and w-bk.

for the three paired opponent-response systems to be selectively stimulated, there must, of course, be more than one substance available for photochemical mediation between the incident light and the neural excitation. Whatever the specific nature of the photosensitive materials, they must form a link in the system of three independent variables,

and hence we have postulated three independent photosensitive materials, which we may call α, β and γ.

Our schematic model now takes the form shown in Fig. 2A or 2B. The three independent photosensitive materials may be contained in discrete retinal units with complex interconnections to the neural response systems, as shown in Fig. 2A, or two or more of these materials may be combined in receptor units having simpler connections to the neural response systems, as diagrammed in Fig. 2B. There is no way of differentiating these models in terms of visual behavior; and however the three photochemicals may be segregated or combined in the retina, and whatever the number of different photoreceptor units, there remain only three independent photosensitive materials, and the theory remains a three-variable, opponent-colors schema.

QUANTIFICATION OF OPPONENTS THEORY

Since our aim is to present this schema in quantitative terms, one of the first questions that has to be asked is this: Is it possible to obtain by psychophysical experiment direct measurements of the spectral distributions of the three basic response variables of the Hering theory?

Measures of Achromatic and Chromatic Responses

It can fairly be assumed that the achromatic, white response is closely connected with the distribution of the brightness quality throughout the visible spectrum, and Fig. 3 therefore shows two functions (which we have measured by a threshold technique) that give the whiteness distribution of an equal energy spectrum for two observers (20). The induced rather than directly stimulated black component of the achro-
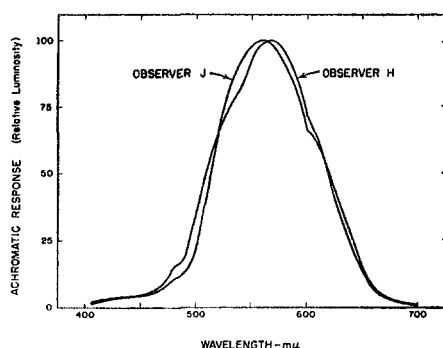
FIG. 3. Whiteness distribution of an equal energy spectrum for two observers.

matic white-black response pair has this same distribution, but of opposite sign, since the strength of the black contrast response is directly related to the magnitude of either the surrounding or the preceding whiteness or brightness.

A method for determining the spectral distributions of the paired chromatic responses is implicit in the opponents theory itself. Since the two members of each hue pair are mutually opponent or exclusive, then a yellow response of given strength should be exactly canceled by a stimulus that, taken alone, elicits the same magnitude of blue response, and a similar relation should hold between red and green responses. Thus a null method, based on the antagonism of the two members of each hue pair, can be used to measure the spectral distributions of the chromatic responses. In brief, a wave length is first selected that evokes, say, a blue hue response. The observer then views, in turn, a series of spectral wave lengths that appear yellowish in hue (yellow-greens, yellow, and yellow-reds). To each of these yellow stimuli just enough of the previously selected blue stimulus is then added exactly to cancel the yellow hue without introducing any blueness. The observer simply reports when the test field appears neither yellow nor blue; the hue remainder that he sees

may be green, neutral, or red, depending on the test wave length. Knowing the energies of the series of spectral yellow stimuli, and having determined experimentally the energy of the blue stimulus of fixed wave length that is required for the hue cancellation in each case, we can now plot the distribution of the relative magnitudes of yellow hue response evoked by the various test wave lengths. The procedure is simply reversed to obtain the distribution of
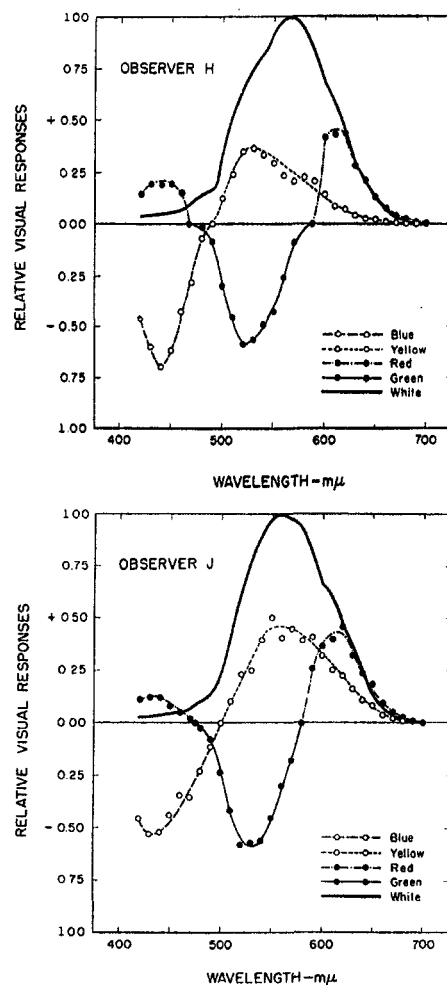

FIG. 4A and 4B. Chromatic and achromatic response functions for equal energy spectrum for two observers.

the blue component of the yellow-blue pair; that is, varying amounts of a fixed wave length of yellow hue are used to cancel the blue hue quality of a series of "blue" test wave lengths. By using a red stimulus of fixed wave length and variable energy to cancel the greens, and a green stimulus to cancel the reds, the spectral distribution of the red-green pair of chromatic responses is similarly determined.

Two sets of paired chromatic response vs. wave length functions that were measured in this way (25), together with the achromatic spectral functions shown in Fig. 3, are plotted in Fig. 4 for an equal energy spectrum. The opponent members of each hue pair have been given arbitrary positive and negative designations, to correspond with their opponent characteristics. Thus the positive values of the red-green function indicate redness, and the negative values greenness. Similarly, the positive values of the yellow-blue function indicate yellowness, and the negative values blueness.

These are the psychophysical functions that represent the spectral distributions of the three independent variables of the Hering opponent-colors theory for these two observers. They are assumed to be directly correlated with the response activity of the visual nervous tissue (retina, optic nerve, and visual centers), and should not be taken as photochemical absorption spectra, about which these data tell us nothing.

Brightness, Hue, and Saturation

The psychophysical opponent-response functions shown in Fig. 4 provide a direct description of the appearance of the spectrum, for these observers, for a neutral condition of bright adaptation and at the moderate level of luminance for which the functions were obtained. Thus, all wave lengths evoke some whiteness as well as hue; the whiteness

and brightness of an equal energy spectrum is relatively small at the two spectral extremes and relatively high at the intermediate wave lengths.

The short wave lengths appear as red-blue hues (violets); there is a narrow band of pure or unique blue where the red-green function is equal to zero; then come the blue-greens, followed by a narrow band of unique green at the wave length where the yellow-blue function is equal to zero; this is followed by the yellow-greens, and then pure yellow occurs at the second intersection of the red-green function with the zero ordinate value; and finally the yellow-red hues appear in the long wave length region (19). A quantitative expression for hue, a "hue coefficient," can be obtained by taking the value of one of the chromatic responses, say, the yellow value at 550 mμ, relative to the total of all chromatic responses at that wave length, in this case, yellow plus green.

The saturation of the color depends on the relative amounts of chromatic and achromatic responses. At the two spectral extremes where the chromatic responses are large relative to the white response, the spectral saturation is high. Where the reverse is true, spectral saturation is low. This can be expressed quantitatively in the form of a "saturation coefficient." To use the same example, the total of the yellow-plus-green values relative to the white plus yellow plus green is relatively low at 550 mμ, and this wave length appears much less saturated than does, say, either 440 mμ or 670 mμ.

Color Mixture

Since color-mixture experiments simply involve matching the three perceived qualities evoked by one stimulus by the proper mixture of three other stimuli, it is possible to determine the color-mixture relations that are inherent in the response curves of Fig. 4 for

any three arbitrarily selected mixture primaries. That is, the red-green value, the yellow-blue value and the white value of the total visual response to any wave length of unit energy are matched by the totals of the three corresponding values for the three mixture primaries when the latter stimuli are combined in the proper ratios. On paper, the color equations for most spectral matches require the admission of negative values for one of the mixture primaries. In actual color-mixture experiments, these negative values are realized by removing one of the mixture primaries from the matching field and adding it to the test stimulus.

To calculate, for example, the amounts of energy required for a color match to a given wave length λ by the mixture of the spectral primaries 460 mμ, 530 mμ and 650 mμ, let a = the energy at 460 mμ, b = the energy at 530 mμ, and c = the energy at 650 mμ. The three equations to be solved for these three unknowns a, b, and c are then:

$$a(r_{460}) + b(r_{530}) + c(r_{650}) = r_\lambda$$
$$a(y_{460}) + b(y_{530}) + c(y_{650}) = y_\lambda$$
$$a(w_{460}) + b(w_{530}) + c(w_{650}) = w_\lambda$$

The values for r (or for $-r$ when the response function is negative, indicating that the hue is green rather than red), for y (or for $-y$ when the response is blue rather than yellow), and for w are then read from the response functions for unit energy for each wave length in question. (See Fig. 4.) The values r_λ, y_λ and w_λ represent the unit energy response values for any spectral wave length for which a color-mixture equation is to be calculated. Solving this set of three equations for the three unknowns a, b, and c, we then have a color-mixture equation of the form

$$a_{460} + b_{530} + c_{650} = 1_\lambda$$

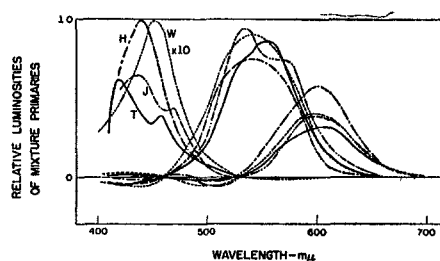This equation, which is expressed in energy units, may be converted to photo-

FIG. 5. Calculated color mixture functions for observers H and J and experimental color mixture functions for observers T and W (24, 47).

metric units in the usual way by multiplying each energy value by the relative luminosity (given by the achromatic response function) at the given wave length.

Color-mixture relations calculated in this manner for wave lengths λ from 420 mμ through 700 mμ from smoothed visual response data for two observers are shown in Fig. 5. The two additional sets of color-mixture functions (for the same three mixture primaries) that are shown for comparison in the figure are the results of actual color-mixture experiments by W. D. Wright and L. C. Thomson (24, 47).

Since the relations between the measured response functions and the color-mixture data are, as we have just seen, known for two individual observers, it is now also possible (by assuming specific spectral loci for the unique hues) to reverse the procedure and derive opponent-response functions from the color-mixture data for Wright and Thomson, or for any other observer whose color-mixture data are available. Since it seems preferable to develop a general theoretical model on the basis of a representative average, rather than a small number of individual observers, we have used for the model chromatic and achromatic response functions derived from the average color-mixture data for the CIE international standard observer (30). These derived functions

are shown in Fig. 6. (The details of the derivation are given in 22.) They are, of course, smoother and more regular than the individual, measured functions, but in other respects they are quite similar (compare Fig. 4).

Photochemical Distributions

The specific set of α, β, and γ photosensitive absorption functions that have been assumed for the theoretical model are shown in Fig. 7. These curves have not been measured, and they have the particular forms shown and the high degree of overlap exhibited because of the specific interrelations that we have postulated a priori between the photochemical events and the neural response activities of the visual opponent mechanisms. Once the photopigments actually present in the human retina have been identified by the biochemists, the visual theorist will have no need to make such a priori postulates, and the specific interrelations required between the identified photosensitive materials and the neural processes underlying the color responses can easily be deduced.
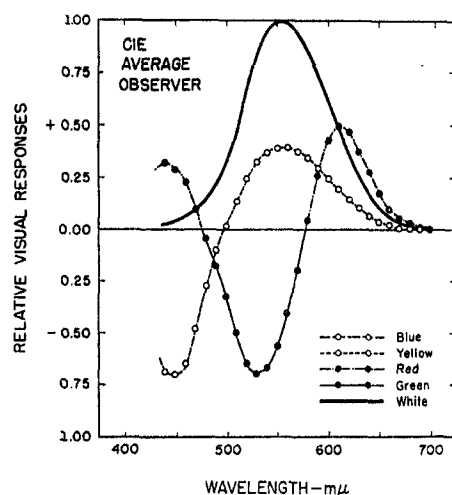

Fig. 6. Theoretical chromatic and achromatic response functions for equal energy spectrum for CIE average observer.
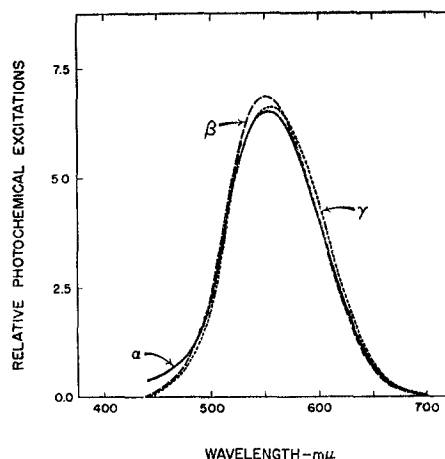

Fig. 7. Spectral distribution curves for assumed photosensitive materials.

As matters now stand, however, the functions shown in Fig. 7 meet the basic demands of the known facts, and any changes in these theoretical absorption functions that will no doubt be required by results of photochemical researches will not importantly affect any of the basic postulates of the theoretical model. The broadness and similarity of shape of all three selective functions that we have assumed are characteristic of all visual pigments so far identified in a variety of animal species (2).

These assumed photopigment distributions do not enter into the consideration of color phenomena, in normal vision, until we come to an examination of some of the phenomena of chromatic adaptation in which selective photochemical bleaching seems to act as one of the important determinants. The other determinants are, of course, the spatial and temporal induction effects in the neural opponent response processes that have been mentioned earlier.

Dependence of Hue and Saturation on Both Wave Length and Luminance

What are the visual phenomena for which the model can account? As we

have already indicated, the measured chromatic and achromatic response functions provide a direct and quantifiable description of the color sensations evoked by any stimulus of specified wave-length composition (23). The achromatic, white function is taken as a direct expression of spectral brightness. Spectral hue, which is determined by the chromatic responses evoked by each wave length, can also be expressed quantitatively as a coefficient value relating the magnitude of response of one chromatic system to the total of all chromatic responses at that wave length. An example of such a hue coefficient function for a moderate level of luminance is shown in Fig. 8. It is clear, from the varying rate of change in the hue coefficient function from one region of the spectrum to the next, that an observer's ability to discriminate among neighboring wave lengths on the basis of hue changes alone will also differ for the different regions of the spectrum. This discriminative capacity is obviously also quantifiable in terms of the amount of wave-length change required to elicit a threshold change of fixed amount in the value of the hue coefficient. With change in the luminance at which the spectrum is presented, these coefficient

WAVELENGTH-mμ

FIG. 8.   Spectral hue coefficients.   Moderate luminance.

functions will be altered, in the sense that the yellow-blue values will increase at the higher luminances, and will be diminished at the lower luminances. This is so because, in accordance with the different energy-vs.-response function postulated for the yellow-blue system as compared with the red-green one, as the excitation level is increased, the yellow and blue spectral responses will be uniformly magnified relative to the red and green ones at the higher levels, and uniformly diminished at the lower levels. Although the exact differential between the two paired systems is not known, under certain circumstances an over-all difference in response magnitudes of approximately 20 per cent seems to occur for a log unit change in luminance. Thus, at some wave length for which, say, the red and yellow responses are equal at a luminance of 10 mL, the yellow will be about 20% greater than the red at 100 mL, and about 20% less at a luminance of only 1 mL. If we assume this 20% differential between y-b and r-g response magnitudes per log unit of luminance change as a reasonable value, and compute the spectral hue coefficients for a range of approximately three log units of luminance variation, then we can specify the amount of hue shift associated with a change in intensity of any wave length. Conversely, we can also specify the wave length changes necessary to maintain a constant hue sensation (constant hue coefficient value) as the luminance is increased or decreased. The latter procedure has been used to obtain the functions shown in Fig. 9, and the curves in the upper part of the figure are functions measured by Purdy in an actual experiment of this sort (38).

These hue phenomena do not involve the achromatic response pair at all, and depend only on the two paired chromatic response systems. Whatever the
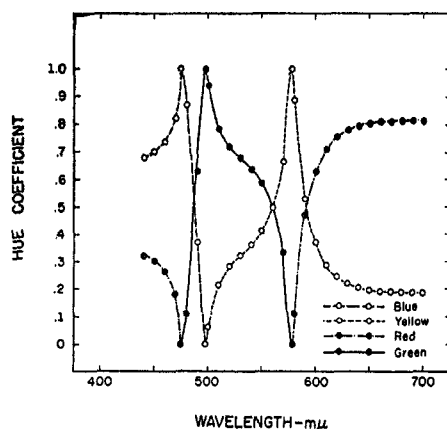
chromatic response to a given stimulus, the perceived color saturation clearly will also depend on the extent to which the white system is simultaneously responding. For any given amount of chromatic response, the color will obviously appear less saturated if there is a large magnitude of white response to dilute the color, and more saturated if the white component of the total response is relatively small. The perceived saturation of the spectrum is also expressed as a quantitative coefficient function. (See Fig. 10.) Here the value taken as the saturation coefficient is the ratio of the total chromatic to the chromatic-plus-white responses at each wave length. The relatively high values at the spectral extremes and the minimal value in the pure yellow region are perfectly consistent both with qualitative reports and with the experimental data on this problem (e.g., 28). Again, as in the hue functions, the rate of change of the saturation coefficient from one spectral region to the next is indicative of a varying discriminative capacity with respect to wave length; and, again, the form of the function as shown applies to a moderate luminance level and varies in a determinable man-
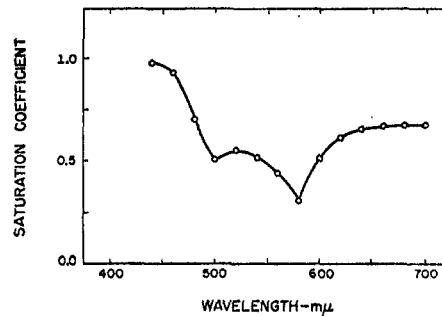

Fig. 10. Spectral saturation coefficients. Moderate luminance.

ner with change in the level of excitation in accordance with the different energy-vs.-response rates of the three independent response systems.

In view of the variations in the hue and saturation functions with change in luminance, we should expect that discrimination functions that depend on changes in these two color attributes, such as discrimination of one wave length from the next in an equal brightness spectrum, would also reflect such a dependence on luminance. Figure 11 shows, in the upper half, a set of wavelength discrimination functions obtained at two luminance levels by Weale (45).
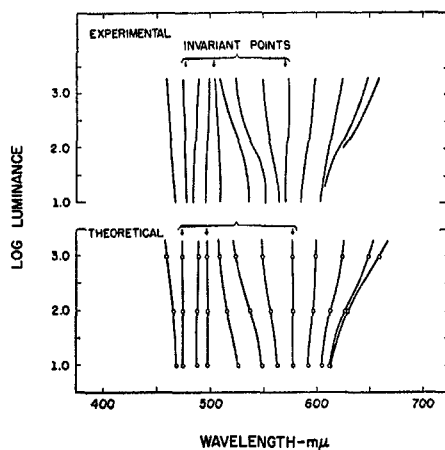

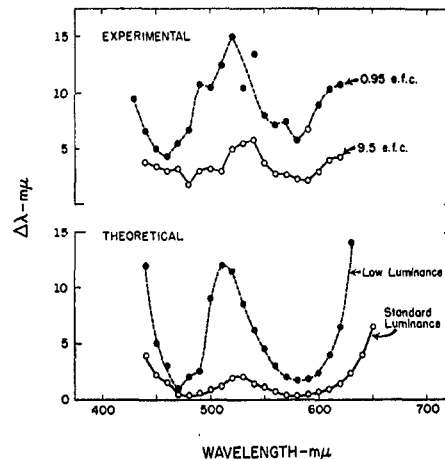Fig. 9. Constant hue contours as measured by Purdy (38) and as predicted by theory.

Fig. 11. Wave-length discrimination functions for two luminance levels as measured by Weale (45) and as predicted by theory.

The higher values of difference threshold obtained at the low luminance level may be explained by a general reduction of discriminative capacity in dim light. The shift of the midspectral maximum toward lower wave lengths, and the relatively greater heightening of the minimum in the yellow region, cannot, however, be attributed to such a generalized reduction in discriminatory capacity. The selectively greater loss in yellow and blue responses at the low-luminance level that is one of the postulates of our model does, however, account for changes of exactly this sort in the form of the function. This is shown by the two theoretical functions computed from pairs of spectral hue and saturation functions that are associated with the two specified luminance levels. Since brightness is kept constant in such experiments, only the hue and saturation variables need be considered in our analysis of these functions (22).

Chromatic Adaptation

The phenomena that we have treated thus far all refer to the individual with normal color vision in a neutral state of adaptation. What of his color perception after the visual system has been exposed for some time to a strongly colored illuminant? For analytical purposes, the simplest situation of this sort is the one in which the eye has been exposed to a large surround field of given color and luminance, and the test stimuli are viewed at the same level of luminance as the surround. Under these circumstances, the three photochemical receptor substances will probably have undergone some selective bleaching, and because of the similar brightness of the surround and test fields, spatial induction effects in the neural response processes will probably be fairly constant. To simplify the treatment for these particular conditions, therefore, we may ignore the constant neural inductions

and consider the photosensitive changes as exercising a controlling influence on the response systems.

We know that under these circumstances the color-mixture data do not change. That is, with uniform chromatic adaptation, any change in the perceived color of one side of a bipartite color-mixture field will also occur on the other side, and to exactly the same extent. Thus a color equation that has been made with the eye adapted to a neutral white light will also be a valid equation when the eye is adapted to a colored illuminant (15). These important constancies of color equations mean that whatever photochemical changes occur with adaptation must occur in a very specific way. That is, the spectral distribution functions representing the three selective photochemicals may be selectively multiplied or reduced by a constant factor, but no one of them can change its form (44, pp. 211–212). In other words, any single substance cannot lose a greater percentage of its absorption at one wave length than it loses at another wave length. Thus, exposure to a colored light can cause any one of the postulated photochemical functions shown in Fig. 7 to be multiplied or divided by a constant amount, but this is the only alteration in the photosensitive functions that is consistent with the fact that color equations are invariant with chromatic adaptation.

The extent to which the three substances are selectively attenuated as a result of exposure to colored light is clearly controlled by the light stimulus itself. That substance which initially absorbs most of the adapting light will suffer the greatest relative bleaching, and the substance which absorbs relatively little of the adapting light will be relatively little affected by it. Thus, by determining their relative absorptions of the adapting light, we can compute the relative changes in the heights

of the photosensitive distribution functions for the three photopigments that we have postulated. Since the excitations of the opponent response systems depend on these photochemical light absorptions (see Fig. 2), we can now also determine the forms and magnitudes of the chromatic and achromatic response functions for the new condition of adaptation. In spite of the close overlap of the photosensitive functions that we have postulated, the "adapted" chromatic response functions determined in this way change in striking fashion relative to the functions for the neutral adaptation condition. The achromatic function changes too, but relatively very little. These theoretically computed adaptation changes are consistent with the kinds of change known to occur in situations of this sort. If the eye that has been adapted to white light is exposed for some time to a saturated red equal in brightness to the white, the normally red end of the spectrum does not become excessively dark, but the amount of redness seen is strongly reduced, and the greens become greatly supersaturated (3, pp. 133–137). Also, the wave length that formerly appeared pure yellow is now strongly greenish, and this is also true for the wave length that formerly appeared pure blue. These changes can be determined from the functions shown in Fig. 12 that have been computed for a given red adaptation, in comparison with the functions for the neutral state that were given in Fig. 6.

From this new set of "adapted" opponent response functions the hue and saturation coefficients and the discrimination data for this new state can also now be determined (26).

These "adapted" response functions are specified, as we said above, for a circumscribed set of conditions for which the photochemical adaptation changes could be taken as primary. As
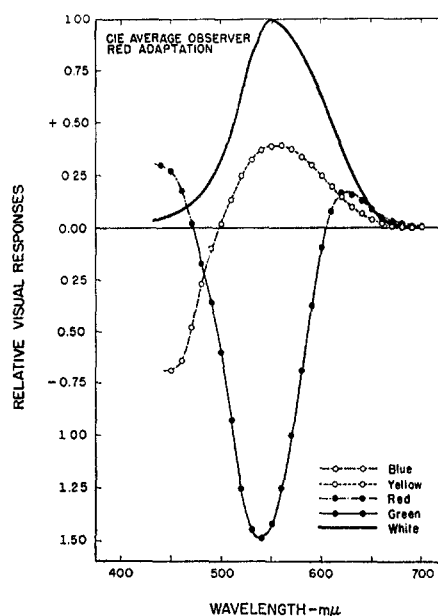

FIG. 12. Chromatic and achromatic visual response functions for red adaptation as predicted by theory.

soon as the relative luminance conditions are altered, however, then the neural inductions enter importantly into the complex picture. For example, if a test stimulus seen within a strongly colored (say, red) surround looks neutral when its luminance is the same as that of the surround, then it will probably appear somewhat reddish at a higher luminance, and the complementary green at a lower luminance (12). (The test stimulus is assumed also to be of predominantly long wave length composition.) In terms of opponent inductions this phenomenon is readily understood. If the red process excited by the red surround induces an opponent green process in the test area, then at an intermediate luminance this green induction is just strong enough to cancel the red-process activity aroused by the test stimulus itself. When the test stimulus is made brighter and the red response to it increases, the unchanged

green induction from the surround becomes inadequate to cancel completely the increased red response to the stronger test stimulus, and the red test hue is now seen. At a much lower luminance of test stimulus, the red process is activated to a much lesser extent, and the green induction from the surround, which is still unchanged in strength, is now sufficient to impart a green hue to the perceived test area. These phenomena are not only consistent with the opponent induction postulate, but they also make it clear why attempts to treat the problem of chromatic adaptation exclusively as a matter of photochemical bleaching are foredoomed to failure (e.g., 1, 33).

Color Anomalies and Color Blindness

When we come to consider individuals who do not have normal color vision we find that their color vision can depart from the normal in two general ways. Their color perceptions may be distorted relative to the normal, or they may exhibit specific color weaknesses or losses. Also, they may show both types of these deviant characteristics at the same time. By distorted color perceptions we mean, for example, the perceptions of the particular type of anomalous individual who has the following characteristics: he sees a distinct orange in the spectral region described normally as pure yellow or nearly so; he needs three stimuli for color mixture; he makes color matches with high precision but uses quite different proportions of the mixture stimuli than does the normal observer. An individual of this type does not seem to have lost any of the efficiency of his neural visual response processes, and it seems reasonable to assume that his color distortions have their basis in the photochemical complex responsible for selective light absorption.

The particular assumptions that we have made concerning the kinds of deviation that the photosensitive materials may exhibit stem from a generalization made by Dartnall (2), on the basis of his researches concerned with the identification of visual photopigments in a variety of lower organisms. Dartnall has found that when the absorption curves of the various visual pigments are plotted as a function of the vibration frequency of the incident light (the reciprocal of the more usual wave-length specification), all the absorption curves have very nearly the same shape, and they can be made to coincide simply by shifting the curves so that they all reach an absorption maximum at the same frequency. In other words, a single template representing amount of absorption as ordinate, against frequency of radiant energy as abscissa, can be used to fit the absorption function of any visual pigment, whatever the locus of its absorption maximum. It seems reasonable to expect that this same generalization will apply to the photosensitive distributions of anomalous individuals with respect to the population of observers with normal color responses. We have consequently assumed that, in congenital abnormalities of the visual system, the normal photopigments can undergo changes that result in a uniform shift of the entire set of photosensitive distribution functions as a group along the frequency scale. These shifts are assumed to occur in either of two directions: toward higher frequencies (shorter wave lengths) resulting in the type of anomalous color vision identified as *protanomaly*, or toward lower frequencies (longer wave lengths) relative to the normal absorption loci, resulting in the second major type of anomalous color vision known as *deuteranomaly*. The amount of these displacements may also vary in different degrees of congenital anomaly.

Since the absorption of light by the photosensitive materials provides the stimulus for the neural chromatic and achromatic response systems, the visual response functions thus controlled by the deviant photosensitive materials will necessarily be altered, too, and in a systematic manner. Examples of theoretically derived anomalous response functions based on these assumptions are given in Fig. 13. The set of functions in the center block are those for the observer with normal photosensitive materials; those in the upper block are for a protanomalous type whose visual pigment absorptions are assumed to be shifted toward the shorter wave lengths by an amount equal to about 15 mμ
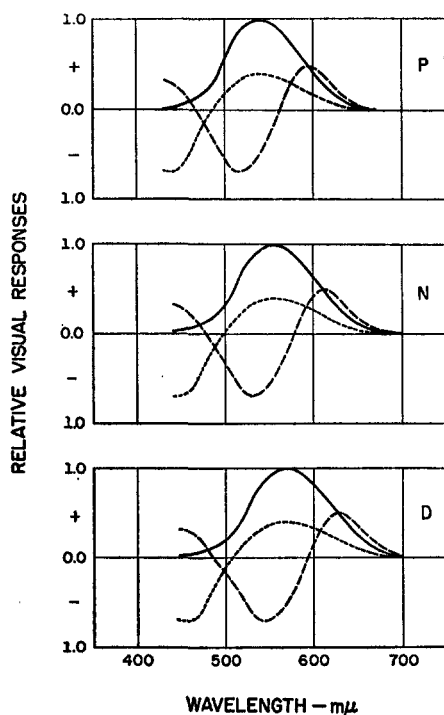

WAVELENGTH — mμ

FIG. 13. Theoretical chromatic and achromatic response functions for equal energy spectrum. For observers with protanomalous, normal, and deuteranomalous photoreceptor systems and with normal strength visual response processes.

from the normal peak of about 550 mμ. This type of individual will have a luminosity function (described by the achromatic, white response function) that peaks at a shorter wave length than the normal and will show considerable loss of luminosity at the red end of the spectrum (48, Ch. 25). The spectral hues will also be altered, with a distinctly reddish yellow occurring where the normal individual sees a unique or pure yellow, whereas the protanomalous observer's pure yellow occurs at a wave length described by the normal as quite greenish. In making color matches, such as a match between 589 mμ on one side of a bipartite field and a mixture of 530 mμ and 670 mμ on the other, this observer will require a much greater proportion of 670 mμ in the mixture than will the average observer with normal color vision (27, 46). This particular match, the Rayleigh equation, is the earliest and best known diagnostic test for anomalous color vision. In this same test, the anomalous individual whose response functions are shown in the lower block in Fig. 13 will deviate from the normal in the opposite way; that is, he will require a much greater proportion of 530 mμ in the mixture for the Rayleigh equation (46). This type of anomalous individual (deuteranomalous) is assumed to have photopigment absorptions that are shifted toward the longer wave lengths, and he will see greenish-yellows where the normal sees yellow, yellows where the normal sees orange, etc. Since the neural response processes of both types of anomalies of this sort are assumed to be operating at the normal efficiency, these individuals will show high precision in making their distorted color matches, and their discriminatory capacities will also be good. As a matter of fact, anomalous individuals of this sort have understandably high confi-
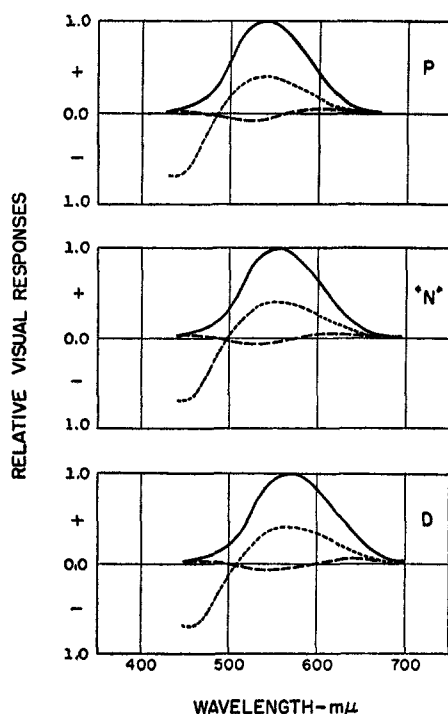
FIG. 14. Theoretical chromatic and achromatic response functions for equal energy spectrum. For observers with protanomalous, normal, and deuteranomalous photoreceptor systems, and with impaired red-green response processes.

dence in their own color capability, and they are extremely resistant toward accepting the results of diagnostic tests which indicate that their visual capacities are deviant from (with the implication of "inferior to") those of the normal population (36, pp. 235–238).

Not all anomalous individuals are as fortunate as the types shown in Fig. 13, however. Many give evidence of real color weakness, in addition to distortions of the kinds already discussed (40). These color-weak individuals seem to have been deprived of some of the efficiency of the neural response processes, particularly of the red-green opponent pair, and their systems may

be represented in terms of the theory by the kinds of response functions given as examples in Fig. 14. The visual pigments of these three types of individuals are taken to be the same as those shown in the preceding figure, respectively, but the red-green paired system is reduced to one-tenth of the normal strength. Such observers have real losses in color discrimination in addition to possible color distortions, and their color matches are imprecise as well as deviant. Individuals with congenitally abnormal color systems are frequently of this general type, and cases of acquired color blindness caused by degenerative disease invariably show this kind of color weakness at some stage in the development of the neural disorder (31).

When the weaknesses become extreme, whether in congenital or acquired disorders, the red-green system seems to be entirely lost to normal function, and a condition of dichromasy, or so-called "color-blindness," results. That is, the visual system becomes a two-variable
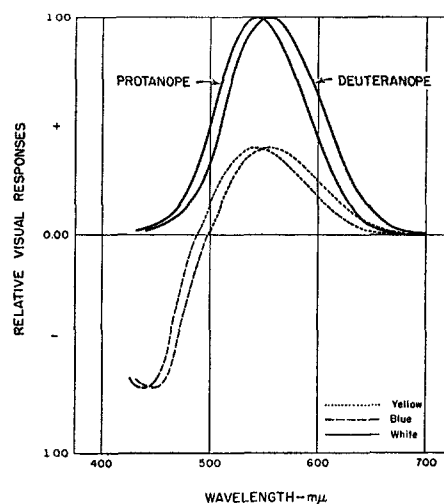

FIG. 15. Theoretical chromatic and achromatic response functions for equal energy spectrum. For observers with nonfunctioning red-green response processes.

one, as shown in Fig. 15. Here the yellow-blue and the white-black neural systems remain intact and functioning, but there is no red-green response function. If the red-green loss occurs without changes in the visual pigments, the remaining yellow-blue and white-black response functions are like those of the normal individual; but, since there is no red-green system, the spectrum is divided into only two hue sections for these individuals. The short wave lengths which normally vary from violet through blue and blue-green to pure green all appear as blue, but of varying saturations, with a neutral region where the normal pure green occurs. Beyond this wave length the remainder of the spectrum appears yellow, in varying saturations, out to the extreme long-wave limit of visibility. The luminosity function is the same as for the observer with normal color vision. Individuals who fit this response pattern would be classified as *deuteranopes* (29). If the visual pigments are altered, so as to produce an absorption shift toward the short wave lengths in addition to the complete red-green neural loss, then the spectrum is again divided into a short-wave blue and a long-wave yellow section, but the neutral region that divides the spectrum into the two major hues occurs at a shorter wave length than for the deuteranopes. The luminosity function is also displaced in this type of dichromasy, as it is for the anomalous individuals with similar photopigment changes, and the type of "color-blind" vision associated with this pattern is called *protanopia* (29).

These two theoretically assumed kinds of deviation from the normal system— i.e., photopigment changes and neural losses or weaknesses of the paired red-green response system—permit us to assemble a systematic picture of the many various manifestations of abnor-

mal red-green vision that defy understanding in terms of any model of the visual system that assumes a one-to-one correspondence between light absorption in the retinal receptors and the resulting color sensations (22, 27).

Defects or losses may also occur in the yellow-blue neural response system, although such defects seem to be much more rare than the red-green defects. Again, these yellow-blue neural losses may take place either with or without changes in the photosensitive materials in the retina. Examples of the remaining red-green and white-black response functions in two types of yellow-blue blindness are given in Fig. 16. In each type of this disorder, the yellow-blue neural response function is missing, and the total gamut of colors for these individuals includes only neutral and reds and greens of various saturations. If there is no simultaneous photopigment disorder, there are two neutral points in the spectrum, one in the region where the normal sees a pure yellow, and an-
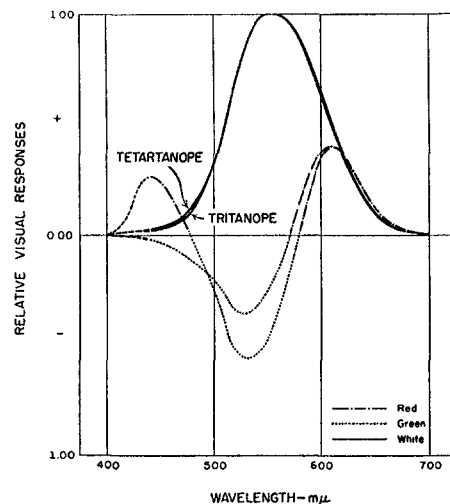

FIG. 16. Theoretical chromatic and achromatic response functions for equal energy spectrum. For observers with nonfunctioning yellow-blue response processes.

other in the region where the normal sees a pure blue. Yellow-blue blindness of this sort is called *tetartanopia*, and only a few cases of it have been reported in the literature (e.g., 34, pp. 68–92). Slightly more common is the second type of yellow-blue blindness, known as *tritanopia* (49), in which not only the neural yellow-blue system is lost, but also the short-wave photopigment seems to be missing. Observers of this type have a neutral point in the normally yellow-green region of the spectrum, but there is no second neutral point, and the green hues extend into the short-wave region that appears violet to the person with normal color vision.

For all these types of deviant color vision, calculation from the theoretical spectral response functions of discrimination curves, color mixture equations, and other psychophysical relations are in good agreement with the experimental data that are available for the various kinds of defective color systems (22, 27).

Opponents-Theory and Neurophysiology

The conceptual model for the opponent-colors theory as originally presented by Hering drew its sharpest criticism on the grounds of being bad physiology. Some of this criticism was based on an erroneous interpretation of Hering's views, an interpretation that incorrectly assigned the opponent processes to the photochemical activities in the retinal cells. Hering's own concept of mutually opponent neural processes, each capable of being activated by external stimulation, was also, however, far ahead of the knowledge of neurophysiology at the time it was proposed (16). But this concept now turns out to be perfectly consistent with the picture of neural function that is only just recently beginning to build up from elec-

trophysiological studies of the visual neural apparatus.

It has become clear that nerves do not simply respond or fail to respond when a stimulus is presented to the appropriate end-organ. Rather, they may respond according to any of a number of quite specific patterns. For example, a nerve fiber may (*a*) discharge at the onset of stimulation and subsequently gradually become quiet; (*b*) discharge at both onset and cessation of stimulation with a quiet period in between; or (*c*) cease any spontaneous activity when first stimulated and during continued stimulation, but respond with a burst of electrical impulses when the stimulus ceases to act (7). The on- and off-phases of discharge are mutually inhibitory processes, they are associated with slow electrical potentials of opposite sign, and they cancel each other when the experimental conditions are so manipulated as to cause both on- and off-discharges to impinge simultaneously on the same ganglion cell (6). In Granit's opinion (6), the evidence from electrophysiology provides a "belated vindication of Hering's view" that the visual system is characterized by mutually opponent neural processes.

The concept of mutual interaction among the various elements of the physiological field is also basic to the theory and is critical to an understanding of both areal effects and simultaneous contrast phenomena. Here again, we find the researches in electrophysiology indicating that individual nerve elements never act independently, and that visual function must be thought of in terms of the integrated action of all the units of the neural visual system (8). Hartline (9) has found that, even in the very simple Limulus eye, the discharge of impulses in any one optic nerve fiber depends not only upon the stimulus to the specific receptor unit

from which that fiber arises but also upon the stimulation over the entire population of mutually interacting elements. Both excitatory and inhibitory interactions of the sort to be expected by theory have actually been demonstrated in the neural responses of the vertebrate visual system by Hartline (8), Kuffler (32), and Granit (6).

The way in which the postulated three independent systems of paired opponent processes (y-b, r-g, w-bk) are differentiated neurally is still a matter for conjecture. Hering thought it was a matter of process specificity, but was willing to use the concept of material, or structural, specificity, which he guessed would be more readily comprehended by most interested readers of his views at the time. Our own theoretical preference at this time is the conjecture that a particular color quality is more probably determined by a particular state of the nervous tissue than by activity of a particular structural element in the nervous network. Thus, we would be inclined to look for a difference between yellow-blue vs. red-green processes, rather than toward isolation of yellow-blue or red-green fibers or nerve cells.

SUMMARY

This paper has presented a summary of our progress to date in providing a quantitative formulation for the Hering opponent-colors theory, and in relating the postulated visual mechanism to specific problems of color sensation, color mixture and color discrimination; to the dependence of these functions on the physical variables of both stimulus wave length and energy level; to their further dependence on adapting and surround stimulation; and to the changes in these functions that occur in various kinds of abnormal color vision. It is

our conclusion that the opponent-colors theory serves as a fruitful working hypothesis by bringing a systematic coherence to the mass of isolated color phenomena that have been reported and subjected to quantitative experiment throughout the years. The physiological concepts basic to the theory are also shown to be consistent with recent findings in neurophysiology.

REFERENCES

1. BREWER, W. L. Fundamental response functions and binocular color matching. *J. Opt. Soc. Amer.*, 1954, **44**, 207-212.
2. DARTNALL, H. J. A. The interpretation of spectral sensitivity curves. *Brit. med. Bull.*, 1953, **9**, 24-30.
3. EVANS, R. M. *An introduction to color.* New York: Wiley, 1948.
4. FARNSWORTH, D. Tritanomalous vision as a threshold function. *Die Farbe*, 1955, **4**, 185-196.
5. GRAHAM, C. H. Vision: III. Some neural correlations. In C. Murchison (Ed.), *A handbook of general experimental psychology.* Worcester: Clark Univer. Press, 1934. Pp. 829-879.
6. GRANIT, R. *Receptors and sensory perception.* New Haven: Yale Univer. Press, 1955.
7. HARTLINE, H. K. The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *Amer. J. Physiol.*, 1938, **121**, 400-415.
8. HARTLINE, H. K. The neural mechanisms of vision. *Harvey Lectures*, 1941-42, **37**, 39-68.
9. HARTLINE, H. K., WAGNER, H. G., & RATLIFF, F. Inhibition in the eye of limulus. *J. gen. Physiol.*, 1956, **39**, 651-673.
10. HARTRIDGE, H. The polychromatic theory. *Documenta Ophthal.*, 1949, **3**, 166-193.
11. HECHT, S. Vision: II. The nature of the photoreceptor process. In C. Murchison (Ed.), *A handbook of general experimental psychology.* Worcester: Clark Univer. Press, 1934. Pp. 704-828.
12. HELSON, H. Fundamental problems in color vision. I. The principle governing changes in hue, saturation, and lightness of non-selective samples in

chromatic illumination. *J. exp. Psychol.*, 1938, **23**, 439–476.

13. HERING, E. *Zur Lehre vom Lichtsinne.* Berlin, 1878.

14. HERING, E. Zur Erklärung der Farbenblindheit aus der Theorie der Gegenfarben. *Lotos, Jb. f. Naturwiss.,* 1880, **1**, 76–107.

15. HERING, E. Ueber Newton's Gesetz der Farbenmischung. *Lotos, Jb. f. Naturwiss.,* 1887, **7**, 177–268.

16. HERING, E. *Zur Theorie der Vorgänge in der lebendigen Substanz.* Prague: 1888. (English translation by F. Welby, in *Brain,* 1897, **20**, 232–258.)

17. HERING, E. *Grundzüge der Lehre vom Lichtsinn.* Berlin: Springer, 1920.

18. HOCHBERG, J. E., TRIEBEL, W., & SEAMAN, G. Color adaptation under conditions of homogeneous visual stimulation (Ganzfeld). *J. exp. Psychol.,* 1951, **41**, 153–159.

19. HURVICH, L. M., & JAMESON, DOROTHEA. The binocular fusion of yellow in relation to color theories. *Science,* 1951, **114**, 199–202.

20. HURVICH, L. M., & JAMESON, DOROTHEA. Spectral sensitivity of the fovea. I. Neutral adaptation. *J. Opt. Soc. Amer.,* 1953, **43**, 485–494.

21. HURVICH, L. M., & JAMESON, DOROTHEA. A quantitative theoretical account of color vision. *Trans. N. Y. Acad. Sci.,* 1955, **18**, 33–38.

22. HURVICH, L. M., & JAMESON, DOROTHEA. Some quantitative aspects of an opponent-colors theory. II. Brightness, saturation, and hue in normal and dichromatic vision. *J. Opt. Soc. Amer.,* 1955, **45**, 602–616.

23. HURVICH, L. M., & JAMESON, DOROTHEA. Some quantitative aspects of an opponent-colors theory. IV. A psychological color specification system. *J. Opt. Soc. Amer.,* 1956, **46**, 416–421.

24. ISHAK, I. G. H. Determination of the tristimulus values of the spectrum for eight Egyptian observers and one British observer. *J. Opt. Soc. Amer.,* 1952, **42**, 844–849.

25. JAMESON, DOROTHEA, & HURVICH, L. M. Some quantitative aspects of an opponent-colors theory. I. Chromatic responses and spectral saturation. *J. Opt. Soc. Amer.,* 1955, **45**, 546–552.

26. JAMESON, DOROTHEA, & HURVICH, L. M. Some quantitative aspects of an op-

ponent-colors theory. III. Changes in brightness, saturation, and hue with chromatic adaptation. *J. Opt. Soc. Amer.,* 1956, **46**, 405–415.

27. JAMESON, DOROTHEA, & HURVICH, L. M. Theoretical analysis of anomalous color vision. *J. Opt. Soc. Amer.,* 1956, **46**, 1075–1089.

28. JONES, L. A., & LOWRY, E. M. Retinal sensibility to saturation differences. *J. Opt. Soc. Amer.,* 1926, **13**, 25–34.

29. JUDD, D. B. Current views on colour blindness. *Documenta Ophthal.,* 1949, **3**, 251–288.

30. JUDD, D. B. Basic correlates of the visual stimulus. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 811–867.

31. KÖLLNER, H. *Die Störungen des Farbensinnes.* Berlin: S. Karger, 1912.

32. KUFFLER, S. W. Discharge patterns and functional organization of mammalian retina. *J. Neurophysiol.,* 1953, **16**, 37–68.

33. MACADAM, D. L. Chromatic adaptation. *J. Opt. Soc. Amer.,* 1956, **46**, 500–513.

34. MÜLLER, G. E. *Darstellung und Erklärung der verschiedenen Typen der Farbenblindheit.* Göttingen: Vandenhoeck and Ruprecht, 1924.

35. PARSONS, J. H. *An introduction to the study of colour vision.* (2nd ed.) Cambridge: Cambridge Univer. Press, 1924.

36. PICKFORD, R. W. *Individual differences in colour vision.* London: Routledge and Kegan Paul, 1951.

37. PIÉRON, H. La dissociation de l'adaptation lumineuse et de l'adaptation chromatique et ses conséquences théoriques. *Année psychol.,* 1939, **40**, 1–14.

38. PURDY, D. M. The Bezold-Brücke phenomenon and contours for constant hue. *Amer. J. Psychol.,* 1937, **49**, 313–315.

39. RIGGS, L. A., RATLIFF, F., CORNSWEET, JANET C., & CORNSWEET, T. N. The disappearance of steadily fixated visual test objects. *J. Opt. Soc. Amer.,* 1953, **43**, 495–501.

40. ROSMANIT, J. *Anleitung zur Feststellung der Farbentüchtigkeit.* Leipzig: Deuticke, 1914.

41. TROLAND, L. T. Apparent brightness: its conditions and properties. *Trans. Illum. Engr. Soc.,* 1916, **11**, 957–966.

42. TROLAND, L. T. *The principles of psychophysiology.* Vol. 2. *Sensation.* New York: D. Van Nostrand, 1930.

43. VON BEZOLD, W. Ueber das Gesetz der Farbenmischung und die physiologischen Grundfarben. *Ann. Phys. u. Chem.,* 1873, **150,** 221–247.

44. VON KRIES, J. Die Gesichtsempfindungen. In W. Nagel (Ed.), *Handbuch der Physiologie des Menschen.* Brunswick: Vieweg, 1905. Pp. 109–282.

45. WEALE, R. A. Hue-discrimination in paracentral parts of the human retina measured at different luminance levels. *J. Physiol.,* 1951, **113,** 115–122.

46. WILLIS, MARION P., & FARNSWORTH, D. Comparative evaluation of anomalo-scopes. *Med. Res. Lab. Rep.* No. 190, 1952, **11,** No. 7, 1–89.

47. WRIGHT, W. D., & PITT, F. H. G. The colour-vision characteristics of two trichromats. *Proc. Phys. Soc.* (London), 1935, **47,** 205–217.

48. WRIGHT, W. D. *Researches on normal and defective colour vision.* St. Louis: Mosby, 1947.

49. WRIGHT, W. D. The characteristics of tritanopia. *J. Opt. Soc. Amer.,* 1952, **42,** 509–521.