# XIV. SIGNAL DETECTION BY HUMAN OBSERVERS*

Prof. J. A. Swets          P. D. Donahue          Louise Iarussi
Prof. D. M. Green                                Susan A. Sewall

## A. COLOR VISION

The "law of additivity of luminances" has long been regarded as a fundamental principle of vision. The standard procedures for the measurement of light intensity and the specification of color depend very directly upon its validity. Recently, some experiments have been reported which bring this principle into question (1). The implications of its deficiency for photometry and colorimetry are now being widely considered (2).

Our own interest in the manner in which luminances combine has to do with the implication of the result for conceptions of the visual mechanism. In the following, we report the results of some preliminary experiments that examine the additivity problem by means of a procedure that seems to us to be more straightforward than those used previously. We relate these results to some data of a different sort obtained from the same subjects, in a hue-discrimination task, and we discuss in a general way the implications of the combined results for models of the process of color vision.

The term "luminance" as used in this connection refers to the effective intensity of a given light for the eye. The measure of the luminance of a component wavelength of a colored light is obtained by weighting the energy at that wavelength by the sensitivity of the eye to that wavelength, or in other words, weighting the energy by its appropriate luminosity. Luminosities can be determined from a standard function that has been adopted by the International Commission on Illumination.

According to the law of additivity of luminances, sometimes called Abney's Law, the luminance of any colored light can be determined by summing over its component wavelengths the products of energy and luminosity. Thus, for example, in color specification it is assumed that the linear sum of the luminances of the three primaries that are adjusted to match the sample is equal to the luminance of the sample.

In the investigations that indicate a failure of additivity, the type of experimental procedure used was suggested by the technique of color specification: the procedures were equivalent to matching a number of samples, spaced over the visible range, by mixing three primaries (sometimes mixing one of the primaries with the sample to control saturation), and then noting where the equation of luminances did not hold. For example, two colors might be adjusted to be identical in hue, saturation, and brightness, but of different composition; in some instances the computed luminances of the two were unequal. Some investigators permitted gross differences in hue or saturation to exist

between the samples that were to be adjusted to equal brightness, and, as might be expected, found the greatest deviations from Abney's Law.

A very typical report from these earlier studies is that the less saturated sample of a pair of samples requires greater luminance than the other if a match in brightness is to be achieved. It was the nature of this result that made it possible to ignore for a time the reported failures of additivity — it was thought that the observers were merely confusing saturation effects with brightness effects. Anticipating our results, we find that the effect of saturation is small; the extent to which two luminances add, in our experiments, depends more heavily on the difference between the wavelengths of the two than on the saturation of the mixture.

The principle feature of our experiments, when compared with previous ones, is that we employed "weak stimuli," stimuli in the "threshold region." The problem posed for the observers was one of detecting signals, rather than one of judging the brightness of a strong signal. Colored lights were projected in the form of circular targets having a diameter of 3° of visual angle, for a duration of 40 msec, on a large background of white light. Five colors were achieved with Kodak Wratten filters Nos. 70, 72B, 73, 74, and 75. These filters have half-peak widths of approximately 30-60 mμ. Their color names, followed by their dominant wavelengths in parentheses, are as follows: Red(678), Orange(606), Yellow(576), Green(538), and Blue(490). It may be noted that the first four of these are spaced along the right leg of the triangle in the ICI chromaticity diagram, whereas the last is positioned midway on the left leg of the triangle.

In a typical experiment, two colors were selected. One of those was presented to a subject a large number of times so that its energy could be adjusted to yield a predetermined detectability. The other was then presented many times, and its energy was adjusted to yield the same detectability. The detectability of the mixture of the two colors was then determined. This experiment was repeated until results were obtained for ten different pairs of colors.

We shall refer to the "single-color" procedure and to the "summation" procedure, to distinguish the two procedures used in a given experiment with a given pair of colors. In both cases, the psychophysical method used was the two-alternative, forced-choice method. With this method, the stimulus is presented in each trial in exactly one of two short, temporal intervals that are defined for the observer by audible clicks. The stimulus is as likely to occur in one interval as in the other. The observer presses a button to indicate in which of the two intervals he believes the signal occurred. A suitable measure of the detectability of the stimulus is the proportion of the correct choices, $P(c)$, made over a series of trials, a measure ranging from 0.5 to 1.0. In our experiments, each proportion is based on 600 trials.

In the single-color procedure, the energy of the stimulus was set to yield a $P(c) \approx 0.70$. (After initial adjustments of stimulus energy were made in order to
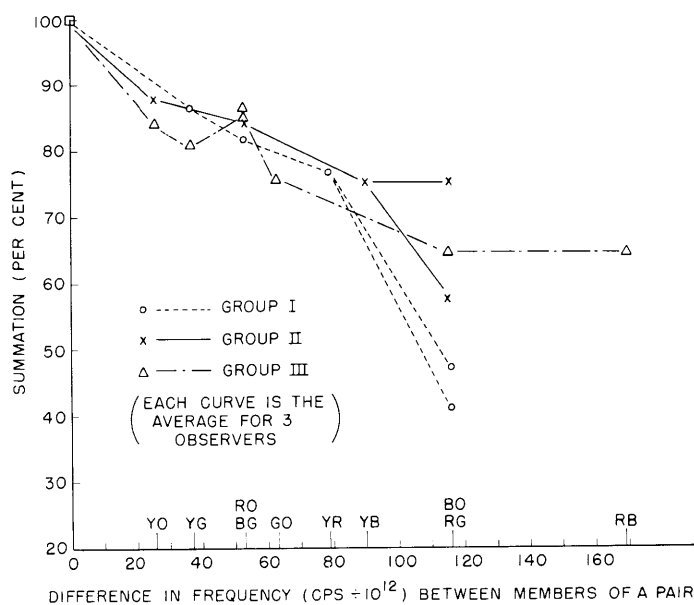
Fig. XIV-1.   Results of summation experiments.

approximate this value, 600 trials were presented to establish the exact value of P(c)
that would be used in further calculations.)   It was determined for each color that a
stimulus of that color having twice the energy of one yielding a P(c) = 0.70 would result
in a P(c) of at least, but just barely, 1.0.   Thus, in the summation procedure, we will
define a complete lack of additivity, or zero summation, as a P(c) of 0.70 for the mix-
ture.   A P(c) = 1.0 associated with the mixture indicates perfect energy summation,
that is, satisfies the law of additivity of luminances.   We interpolate between these
values to obtain an estimate of the per cent of summation that is associated with each
color pair.   These calculations are made individually for the data obtained for each
observer, but in the presentation of the results, the final figure, that is, the per cent
of summation, is the average for three observers.

We have plotted the per cent of summation as a function of the difference in frequency
between the two colors used in a single experiment.   The results for three groups of
three observers are shown in Fig. XIV-1.   Listed along the abscissa at the appropriate
frequency difference are the 10 pairs of colors used, indicated by the first letter of
their names.   It can be seen that the summation is nowhere perfect when a difference
in frequency exists between the two components of the mixture.   (The data point at zero
difference in frequency represents the fact that doubling the energy required for
P(c) = 0.70 for any of the colors results in a P(c) of approximately 1.0.)   Furthermore,
and more important, the per cent of summation that is observed drops regularly as the
difference in frequency of the components increases.   The picture changes little if we
use difference in wavelength, rather than difference in frequency, as the scale.   In this

case, RO moves to the right of GO, YB and YR are reversed in position, and RG moves to the right of BO.

We may note that the saturation of the mixture has little or no effect.  Since four of the colors (all but Blue) are located on one leg of the color triangle, their mixtures are not appreciably less saturated than are the individual components.  The trend is no less definite if we eliminate from consideration those pairs containing the Blue as one member.  Furthermore, the pairs RO and BG, which have the same frequency difference but differ in saturation, yield almost the same per cent of summation.  Looking now at the pairs BO and RG, we see that they do not yield the same result precisely, but the lower point for both Groups I and II corresponds to the RG pair and not to the BO pair — the mixture with the greater saturation shows less summation in this case.

After these experiments were completed, we conducted some others with a slight variation in procedure as a check on the results.  The observers in Group III, the group still in our employ at the time, served in this experiment.  In this instance, we adjusted the energies of the colors in the single-color phase of the experiments to produce a $P(c) \approx 0.90$.  We then halved their energies in the mixture, i.e., in the summation procedure.  In this case, perfect energy summation consists of attaining the same value of $P(c)$, i.e., 0.90, from the mixture. A complete lack of summation would be indicated by a decrement in $P(c)$ of 0.40, which would correspond to chance performance under the two-alternative, forced-choice method.

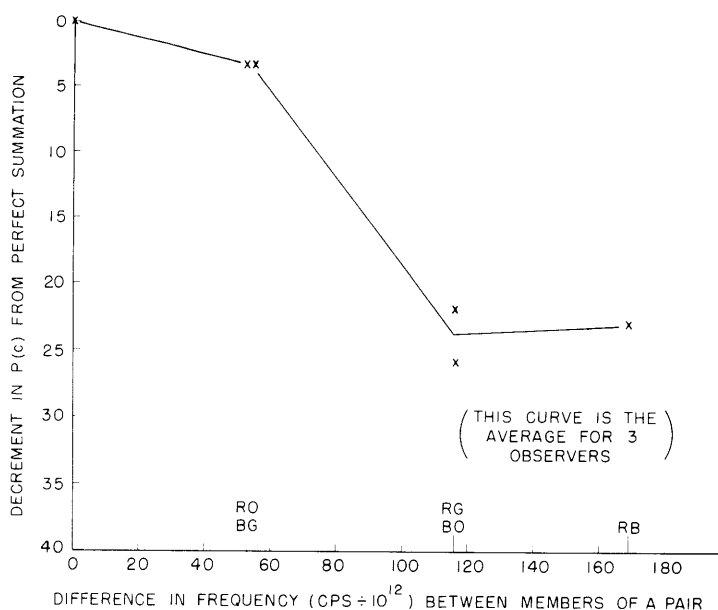The decrement in the value of $P(c)$ associated with the mixture, from 0.90, is



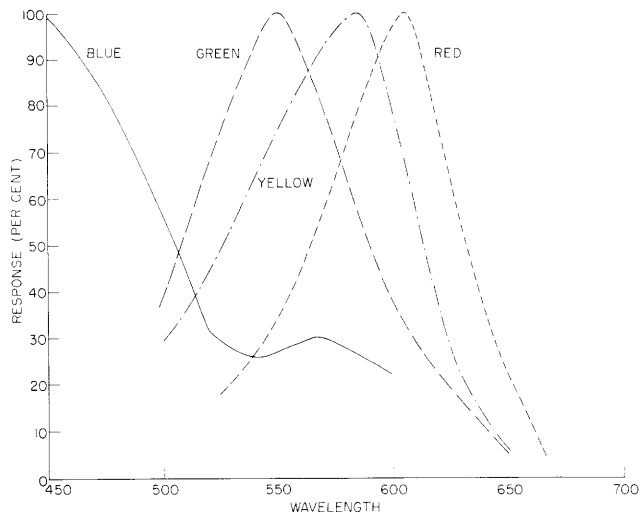Fig. XIV-2.  Results of a second summation experiment.

Fig. XIV-3. Sample of Granit's modulators.

plotted against the difference in frequency in Fig. XIV-2. We see that the amount of summation is less for greater differences in frequency, which supports the earlier result.

Let us assume, then, that the extent to which luminances add is a function of their difference in frequency. As a result, what changes must be made in our conception of the process of color vision? Perhaps no general changes in conception are required. If so, it is difficult to see how the law of additivity of luminances would be consistent with our conception.

Recall that the prominent theories of color vision postulate a number — sometimes three, sometimes four or more — of different types of color receptors. These may be thought of as sensory filters arranged along the frequency axis. The filters can be represented by the "basic sensation curves" that Helmholtz, König, and Hecht inferred from psychophysical data, or by the "modulator" curves that Granit (3) obtained from lower animal forms by electrophysiological techniques. A sample of the latter is reproduced in Fig. XIV-3.

For the purpose of discussion, let us assume an array of sensory filters having the same width and shape. First consider two monochromatic stimuli, each located at the peak of a sensory filter, and ignore, for the moment, the presence of the other sensory filters. A reasonable extension of the general filter models that have been proposed would lead to the prediction that the detectability of a mixture of the two colors will depend upon the difference in frequency between the two, that is, upon the overlap of the two sensory filters in question. In the extended model, the greater the overlap, the greater will be the positive influence of the energy of one color upon the sensory response that is mediated primarily by the other filter. That is, Filter 1 will respond fully to Signal 1, but will also respond to some extent, depending on the amount of overlap, to Signal 2. Similarly, Filter 2 will respond fully to Signal 2, and to some extent to Signal 1. Thus, the outputs of the two sensory filters will have a common factor; the magnitude of

131

this common factor will vary with the amount of overlap of the filters. Furthermore, it seems reasonable that the greater the overlap, the less will be the total noise power in the relevant band — having reference here to both experimenter-produced and observer-produced noise.

This analysis follows one developed by Green for the analogous problem in audition (4). Let us consider another psychophysical task that bears on this question of the dependence between the outputs of different sensory filters, and a related model developed by Tanner (5), another colleague working in audition. We shall consider Tanner's model in more detail, and then return to consider some quantitative aspects of Green's model. Unfortunately, for reasons we shall discuss, we may not be able to put to use in the visual case the quantitative properties of Green's model. Both of these models, it might be noted, are logical extensions of a general theory of signal detectability that is based upon statistical-decision theory. As such, they do not depend upon the validity of any hypothesis about the visual mechanism, but they can be interpreted in terms of a visual filtering process.

The related psychophysical task of which we speak involves what we shall term the recognition procedure. In these experiments, the detectabilities of two colored lights are determined individually as before. Then, under the recognition procedure, one or the other of the colors is presented on a trial containing a single observation interval, and the observer indicates by means of his two buttons which color he believes was presented. The same ten pairs of colors that were used in the summation experiments were used in successive recognition experiments, and again 600 trials are presented to define the detectability of each stimulus and the recognizability (identifiability) of the members of the pair. Actually, these experiments were interleaved with the summation experiments described above.

It is necessary at this point to introduce another measure of detectability, the measure d'. This measure is one suggested by the theory of signal detectability that is based on statistical-decision theory. Its definition has been presented elsewhere (6). Briefly, it is assumed that the process of detection, or of recognition, is a choice between two random variables that are normally distributed and of equal variance. In the detection problem, one of these is associated with noise alone, and the other with signal plus noise. The index d' is defined as the difference between the means of the two distributions, divided by the standard deviation. The noise distribution can be normalized so that its mean is zero and its standard deviation is unity, and then the parameter d' becomes the mean of the signal-plus-noise distribution. It will be intuitively clear that d' is monotonically related to signal strength, and that it is also monotonically related, in a specifiable manner, to the proportion of correct choices in a forced-choice experiment. In the recognition problem, the distribution associated with Signal 1 is assumed to lie along one dimension with a mean equal to $d_1'$; the distribution
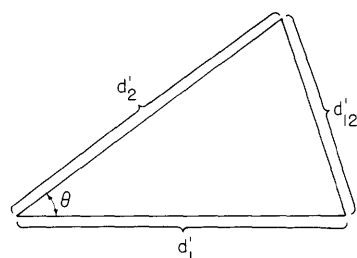
$$(d'_{12})^2 = (d'_1)^2 + (d'_2)^2 - 2d'_1 d'_2 \cos\theta$$

$$\cos\theta = \frac{(d'_1)^2 + (d'_2)^2 - (d'_{12})^2}{2d'_1 d'_2}$$

Fig. XIV-4. The recognition space.

arising from Signal 2 is regarded as lying along a second dimension with a mean equal to $d'_2$. We then assume that the distributions associated with Signals 1 and 2 lie along a third axis that joins the means of the two distributions, and that the means of the two distributions are separated by the quantity $d'_{12}$. This quantity, $d'_{12}$, can be estimated from the proportion of correct responses in a recognition task.

In the experiments under discussion, we obtain two empirical indices of detectability, $d'_1$ and $d'_2$, and an index of recognizability, $d'_{12}$, according to the procedures outlined above. It is assumed that these can be represented in a space as in Fig. XIV-4. This assumption is based on the following line of reasoning. We have two sensory filters, Filter X centered at the frequency of Signal 1 and Filter Y centered at the frequency of Signal 2, with outputs x and y, respectively. We consider a geometrical representation in which x and y are the coordinates of a plane. In the single-color procedure, the observer makes either an observation x or an observation y; in the recognition procedure he makes an observation xy. In the single-color procedure, with a given signal strength, the expected observations are 0 if noise alone is present, $\bar{x}$ if Signal 1 and noise are present, and $\bar{y}$ if Signal 2 and noise are present. In the recognition procedure, in which the observation xy is made, the expected observations are $(\bar{x}, k\bar{x})$ if Signal 1 and noise are present, and $(k\bar{y}, \bar{y})$ if Signal 2 and noise are present, in which the constant k represents the effect of Signal 1 on the output of Filter Y and the effect of Signal 2 on the output of Filter X. The points $(0, 0)$, $(\bar{x}, k\bar{x})$, and $(k\bar{y}, \bar{y})$ define a triangle in the plane. The triangle has an angle $\theta$ at the origin that is bisected by a 45° line; in Fig. XIV-4 the triangle has been rotated slightly.

From a knowledge of the three empirical indices $(d'_1, d'_2, \text{ and } d'_{12})$, we can calculate the angle $\theta$, by using the equations shown in the figure. We shall define the cosine of $\theta$ to be a measure of the dependence between the responses of the sensory filters mediating Signals 1 and 2. Since the signal energies are adjusted so that $d'_1 = d'_2$, the first equation in Fig. XIV-4 can be rewritten as $d'_{12} = \sqrt{2}d'_1 \sqrt{1 - \cos\theta}$. Let us consider the two extreme results. If the responses of the two sensory filters that mediate two colors are completely dependent, that is, if the filters overlap completely, then $\theta = 0$, $\cos\theta = 1.0$, and $d'_{12} = 0$. If the responses of the two filters are completely independent, that is, if the filters do not overlap at all, then $\theta = 90°$, $\cos\theta = 0$, and $d'_{12} = \sqrt{2}d'_1$. Thus, the same reasoning presented above — namely, that there will be a common factor (of varying magnitude) in the outputs of the two sensory filters that are centered,

133

respectively, at the two colors used in the experiment — leads to the prediction that the recognition accuracy will increase with increases in the frequency separation between the two colors.

The experimental results are shown in Fig. XIV-5. In this figure, cos θ is plotted against the difference in frequency between the two members of a pair of colors. Except for a definite reversal at the small differences in frequency, the results exhibit the downward trend that is predicted from the model. The value of cos θ decreases from approximately 1.0 at a small separation, to a value greater than 0 at the largest separation.

The reversal observed at small differences in frequency, though less pronounced there, can also be seen in the summation data shown in Fig. XIV-1. We do not regard this result as sufficient cause to reject altogether the model we have presented. However, if we accept the data, the model must be modified in some way. At this stage of the research we prefer to consider a particular modification of the model that can be described as follows. In presenting the argument we have used the difference in frequency between the two members of a pair of colors as the independent variable. It will be apparent that the points on this scale of differences can be ordered in various ways by various monotonic transformations of the frequency scale, and, also, that our presentation assumed a linear frequency scale. There is, however, no a priori reason to believe that the observer's basic scale is a linear frequency scale. We may note that the high points (cos θ ≈ 1.0) correspond to the pairs RO and BG, that is, to pairs at the
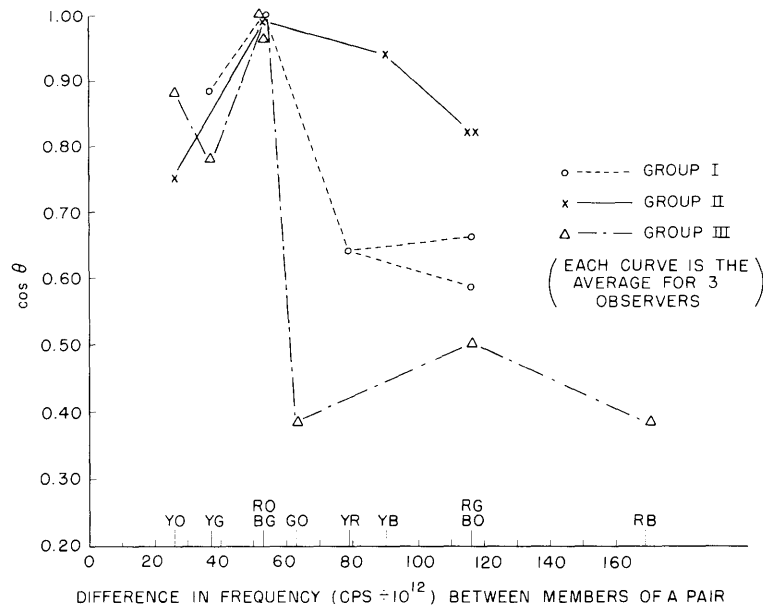


Fig. XIV-5.  Results of recognition experiments.

ends of the spectrum. The lower points to the left of these are obtained with the pairs YO and YG, that is, with colors in the middle of the spectrum. The data we have obtained are therefore consistent with an underlying scale of frequency that is stretched in the middle or squeezed at the ends. Such a scale would result in a difference scale on which YO and YG are found to the right of RO and BG; furthermore, the low point at GO would be moved to the right of the points at YR and YB. An alternative interpretation of the data is that the working assumption of sensory filters of equal widths and of the same shape is not justified — that, in particular, the sensory filters in the middle of the spectrum are narrower than those at the ends. These interpretations of the data, or possible modifications of the model, are not distinguishable on the basis of the data. Either is consistent with the data from a variety of psychophysical experiments that can be found in the literature.

We have said that we would return to Green's model of the process of summation in order to consider its quantitative aspects, and we have also said that it is related to Tanner's model of the recognition process. Green's model leads to the prediction that, for widely separated frequencies affecting independent sensory filters, the value of $d'$ for the linear combination is the square root of the sum of the squares of the values of $d'$ that are associated with the individual frequencies. For the case of equal values of $d'$ associated with the individual frequencies, the predicted detectability of the compound signal of n frequencies, $d'_n$, is equal to $\sqrt{n}d'_1$. This prediction reflects the familiar result in statistics that the standard deviation of a sampling distribution decreases as the square root of the number of samples. The prediction for the sum of two frequencies that have equal values of $d'$, and that are mediated by overlapping sensory filters, reduces to $d'_{12} = \sqrt{2}d'_1 \sqrt{1 + k}$ where k represents the extent of dependence between the outputs of the two filters. For no overlap, $k = 0$ and $d'_{12} = \sqrt{2}d'_1$. For complete overlap, $k = 1$ and $d'_{12} = 2d'_1$.

Thus, in the models for summation and recognition based on statistical decision theory, the predicted value of $d'_{12}$ (where this notation is used to indicate both the index of summation and the index of recognition) is equal to the vector sum and vector difference, respectively, of $d'_1$ and $d'_2$. In his auditory study, Green used only widely separated frequencies such that $k \approx 0$; he refrained from making quantitative predictions for small frequency differences ($k > 0$), since the relationship between k and frequency separation, a relationship which depends on the shape of the filter, is not known. In principle, an empirical determination of this relationship is possible. One might also determine the value of $\cos \theta$ by the recognition procedure and relate this value in the recognition equation, namely, $d'_{12} = \sqrt{2}d'_1 \sqrt{1 - \cos \theta}$, to the value k in the summation equation, $d'_{12} = \sqrt{2}d'_1 \sqrt{1 + k}$.

It is unfortunate that in spite of the fact that the quantitative predictions of the summation model have been confirmed in auditory studies, with widely separated

frequencies, we are unable to apply them in the case of vision. This is true because the model requires that the value of d', for a single frequency, be proportional to signal energy. (Note that the prediction of the model for the case of complete overlap of filters, where $k = 1$, reduces to $d'_{12} = 2d'_1$.) It has been shown, in the case of auditory signals, that d' is approximately proportional to signal energy. However, data that we have collected on white and colored lights show d' to vary as approximately the square of signal energy.

Let us summarize and evaluate our results. We have the long-standing, but recently questioned, assumption that the luminances of different wavelengths add in a linear fashion. Models of color vision that can be interpreted as filter models also have been in existence for some time. An extension of the filter model leads to the prediction that the degree of summation of two frequencies will decrease with an increase in the separation of the two frequencies. The assumption made is that there will be a smaller effect of the energy at one frequency upon the output of the filter that mediates the other frequency as the separation between the two frequencies increases. The same reasoning leads to the prediction that the recognizability of two frequencies will increase with increasing frequency separation. Our experimental results are consistent with these predictions.

As we have mentioned, we have available only the results of preliminary experiments. Furthermore, the model we discussed is developed on only a very general level, and there remain some problems of quantification. We do not pretend to have a detailed specification of the process of color vision. We submit merely that the results we have obtained, if confirmed by further experimentation, are relevant to an effort to construct such a detailed model. We would also consider the possibility that the psychophysical techniques and the models described above, if paired with more adequate experimental apparatus, are capable of determining more accurately than techniques previously available the spectral sensitivities of different types of receptors in the human eye.

J. A. Swets

## References

1. For a list of references to these experiments, see L. M. Hurvich and D. Jameson, Color vision, Ann. Rev. Psych. 11, 101 (1960).

2. D. B. Judd, Radical changes in photometry and colorimetry foreshadowed by CIE actions in Zurich, J. Opt. Soc. Am. 45, 897-898 (1955).

3. R. Granit, The retinal mechanism of color reception, J. Opt. Soc. Am. 31, 575 (1941).

4. D. M. Green, Detection of multiple component signals in noise, J. Acoust. Soc. Am. 30, 904-911 (1958).

5. W. P. Tanner, Jr., A theory of recognition, J. Acoust. Soc. Am. 28, 882-888 (1956).

6. W. P. Tanner, Jr. and J. A. Swets, A decision-making theory of visual detection, Psychol. Rev. 61, 401-409 (1954).