Comparison of Four Methods of Heterochromatic Photometry

GUNTHER WAGNER
Fernseh Darmstadt, West Germany

AND

ROBERT M. BOYNTON
Center for Visual Science, University of Rochester, Rochester, New York 14627
(Received 17 March 1972)

Four methods of heterochromatic photometry were employed, using the same four observers in each case. These were (i) two types of direct heterochromatic photometry (direct comparison with white, and step by step), (ii) flicker photometry, and (iii) the minimally distinct-border method (MDB). The MDB method is shown to yield results that are linear and obey Abney's law. Flicker and MDB methods generate relative luminous-efficiency functions that agree well with each other and also with the CIE standard observer as modified by Judd; the methods of direct heterochromatic photometry yield data that agree fairly well with each other, whereas they differ greatly from the data obtained by flicker or MDB. Luminous efficiency as measured by the direct methods seems to receive a contribution from two sources, (a) achromatic signals of the photopic visual system, which exclusively determine the MDB setting, and (b) chromatic signals of the visual system, which produce extra brightness, the amount of which is related to the saturation of the stimulus used.

INDEX HEADINGS: Vision; Color; Photometry.

The central problem in heterochromatic photometry has been to determine whether two lights of different chromaticity appear equally bright. The basic experimental procedure seems simple enough: The radiance of one of two lights is adjusted until an observer decides that a state of brightness equality has been achieved between them. But this seemingly simple procedure conceals some intricate experimental and conceptual problems.

EXPERIMENTAL PROBLEMS

Direct comparisons of two lights for brightness can be made easily and reliably only if the chromaticities are not too different. As the difference of chromaticity between them is increased, the method of direct comparison becomes increasingly less precise. Although an observer can easily tell when one of the two lights is distinctly brighter or dimmer than the other, there is between these conditions a rather extended range wherein such a decision cannot easily be made. Moreover, the equal-brightness criterion is not stable with time, and differences between observers are large.[1] We would, of course, prefer to find just a single point somewhere in this uncertain range that could represent the condition of equal brightness.

These problems long ago stimulated search for criteria that would be more reliable, and from which one could infer—rather than judge directly—the equality of brightnesses. In principle, any index of visual performance could be used for this purpose, so long as it is monotonically related to radiance; for example, two criterion responses that have been explored are critical flicker frequency[2] and visual acuity.[3] The adoption of various criteria leads to diverse results, meaning that no two lights of different color can be equal in all their possible visual effects at photopic levels.

CONCEPTUAL PROBLEMS

It was of practical importance that the concept luminance be defined so that it would fit into the system of concepts and measures that are useful in physics and applied technology. Thus, in addition to finding a method that was reliable, it was necessary that values of luminance be additive, so that they would conform to the rule of linearity. The criterion adopted, upon the recommendation of Ives,[4] was that of flicker photometry, a method which produced results that proved to be linear within limits, under certain conditions.

The problem was bypassed by the CIE, which defined luminance in accord with the requirement of linearity by use of the equation

$$L=K\int L_{e\lambda}V_{\lambda}\,d\lambda, \quad (1)$$

with V_{λ} being the luminous efficiency function of a standard eye, based largely upon flicker photometry.

In 1932, Guild[5] wrote, concerning the relation between heterochromatic brightness matching and flicker photometry[6]:

"The science of photometry consists of the study of the relationships between all of the properties of light sources and other material objects which influence brightness as measured by a human observer. For practical purposes, such relations are expressed in various forms, and various experimental methods are used to facilitate the rapid and precise evaluation of photometric quantities in routine practice [especially, these days, photoelectric photometers]. For our present purpose, however, we must eliminate from consideration all indirect methods, such, for example, as the flicker method, which can be used only under special conditions for which it is known, from previous experiment, that the results agree with those obtained by one fundamental method [brightness matching] accepted as defining the meaning of the measurements. When, for reasons of practical convenience, we employ any other method to evaluate such quantities, we do so on the presumption, which we must justify by suitable experiments, that the method does in fact perform the valuation in accordance with this meaning, at any rate with sufficient accuracy for our purposes."

Brightness comparison is thus accepted by Guild, and certainly by many others as well, as the fundamental operation upon which a valid heterochromatic comparison should rest. Unfortunately, the trouble with brightness, in addition to the reliability problems already mentioned, is its intrinsic nonadditivity, the failure of Abney's law. In some cases, the failure may be by as much as 50%. In general, where there is a cancellation of hue, there will also be a partial cancellation of brightness. To take a specific example; if yellow and blue fields, set for equal brightness, are added together by optical superposition, the mixture field will be much dimmer and very much less saturated than either of the component fields adjusted to twice its original luminance.

Under practical conditions, such additivity failures are usually minor, and the concept of luminance, as defined by the CIE, usually works out all right in practice and the errors that are made about the relative brightnesses of such things as fluorescent vs tungsten lamps are not serious. Yet there is something unhappy about a system that works only by definition.

Why is the method of brightness comparison intuitively so attractive that it is regarded as the fundamental one? Perhaps it is because the two fields to be compared can be viewed simultaneously, side by side. Is there another method for putting two fields side by side, which could be used instead, and which, with luck, might satisfy the criterion of linearity? There is such a method, that of the minimally distinct border (MDB), first described by Boynton and Kaiser[7] in 1968. Rather than compare the juxtaposed fields for brightness, the subject attends to the boundary between them, adjusting the radiance of one of the fields until the boundary appears to be minimally distinct. This turns out to be relatively easy to do. Most importantly for the present discussion, the results to date conform to the criteria of linearity; more evidence for linearity will be presented here.[7]

We therefore propose that the MDB method is superior to the method of brightness comparison and that the latter should be abandoned as a basic criterion. The MDB would not, however, be useful if it generated a luminosity function having no relation to the standard one. In this paper we show that the spectral sensitivity of the eye assessed by the MDB is very similar to that provided by flicker photometry, and that both agree well with the CIE standard observer.

ELIMINATING THE REFERENCE STIMULUS

The assignment of a particular numerical value to a light is equivalent to subdividing the set of all lights into classes of equal value; that is, all members of one particular class must have the same number assigned to them. A class is represented by any one member of this class; moreover, this representative may be replaced by any other member of the same class. A partitioning of a set into classes is possible if, and only if, there exists a relation of equivalence between any two members of the set.

An equivalence relation obeys the properties of reflexivity, symmetry, and transitivity. In the MDB case, the equivalence relation corresponds experimentally with the adjustment of the radiance of one field with respect to its carefully juxtaposed neighbor, to produce a MDB between them. We should distinguish between two cases. In the first, two lights, A and B, are both compared against a variable reference of known chromaticity. For example, lights A and B are said to be equivalent, if each forms a MDB with the same radiance of a reference white (W). The equivalence relations are then: A constitutes a minimum border with W; B constitutes a minimum border with W. The results hold only for the particular reference white used in the experiment, and in general such results are restricted to the particular reference stimulus chosen, unless greater generality can be proved.

To remove this restriction on the reference stimulus, we must check to see whether two lights A and B that have a MDB with any third reference color also yield a MDB when directly compared with each other. If such transitivity is verified, the equivalence relation reduces to "A constitutes a minimally distinct border with B." The advantage now is that a photometric system based upon this particular form of equivalence relation is not only quite generally valid, independent of the chromaticity of the reference light, but in fact does not require any reference light at all. It is this feature that simplifies the procedure of measurements and allows a practicable operation. It is evident that the operation is reflexive; Kaiser *et al.*[8] have already tested for symmetry and transitivity and have obtained confirmatory results.

APPARATUS

For our experiments, we used a five-channel optical stimulator, by means of which we constructed visual fields of various forms and color combinations. One channel is shown in Fig. 1. The five channels received their light from a dc-supplied 900-W xenon-arc lamp. Wavelength and radiance of the light in each of the five channels was controlled separately and independently by circular neutral wedges and by continuous interference filters. If white light was desired, an interference filter could be removed easily. In every channel there was placed, at a conjugate focus, a field stop of appropri-

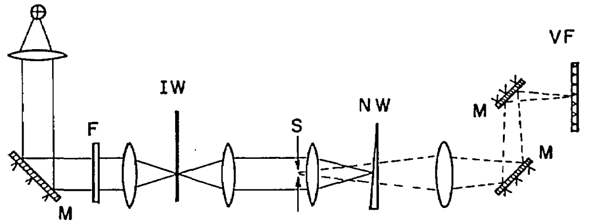

FIG. 1. Optical schematic of one channel. F: heat-absorbing filter; IW: interference wedge; S: visual field stop; NW: neutral density wedge; M: mirror; VF: visual field. V is a magnesium carbonate block; the image formed upon it is viewed through another mirror, not shown.

ate form, which was projected onto a $MgCO_3$ surface, thus providing the contours of visual-field components.

The spatial position of the partial fields could be controlled from the observer's seat in order to maintain satisfactory borders; minor adjustments were frequently required during an experimental session. Observations were made through a 5-mm artificial pupil; an achromatizing lens was used to correct for the chromatic aberration of the eye. The observer controlled the angular position of the neutral wedges and thus adjusted the radiance in the five channels until the criterion under consideration was fulfilled.

In the experiments to be reported here, only three of the five channels were used: One provided white light; the other two provided spectral lights. A bipartite circular photometric field, subtending a diameter of 1°40′, was presented in dark surround. It was viewed continuously, except in the minimum-flicker condition. At the beginning of the experimental sessions, the observer was given sufficient time to adapt to the luminance level prevailing during the experiment.

THE MDB CONCEPT

Given two carefully juxtaposed half-fields, they are seen as a circle divided by a sharp line. Suppose, for example, that the wavelength of one-half of the field is 530 nm, seen as green at some fixed, intermediate, photopic level of radiance, whereas the radiance of the juxtaposed field (say, red at 650 nm) is continuously variable. The border between the two fields is seen as more or less vivid or distinct, depending upon the relative radiances of the two fields.

For very low radiance of the red field, it is seen as dark and the border between it and the green field appears clearly distinct. Increasing the radiance of the red field then decreases the distinctness of the border (this should not be confused with blurring, or defocusing). The subjective brightness difference between the two fields also becomes less marked. There is no radiance level at which the border between the red and green fields vanishes; however, for some radiance there is a minimum distinctness of the border separating the fields. With further increases of the radiance of the red field, the distinctness of the border again becomes greater. The radiance at which minimum border distinctness occurs specifies the result for the MDB criterion. The MDB criterion does not usually result in two equally bright half-fields; generally, the more saturated member of the pair appears brighter at the MDB setting.

MDB EXPERIMENTS TO TEST ADDITIVITY

Procedure

Two spectral lights, λ_1 and λ_2, were adjusted so that their radiances, taken one at a time, fulfilled the MDB criterion when juxtaposed with a constant reference white of 80 td. The relative radiances required to do this are in each case specified as 100 (%). Then we presented some proportion α of λ_1, upon which we superposed λ_2. The observer was instructed to adjust the amount β of λ_2, so that $\alpha+\beta$ yielded a MDB with the reference white. α was varied in steps of 10 from 10 to 90, and the corresponding β was determined each time. During the same experiment, β was also fixed at various values and the corresponding α was determined. Each combination of two spectral lights thus provided 18 pairs of values, α, β. If we plot α vs β, then if the values are additive, $\alpha+\beta=100$, the points ideally should lie on a straight line that connects the points $\alpha=0$, $\beta=100$ and $\beta=100$, $\alpha=0$. Observance of additivity was tested thoroughly in this way by one observer (GW) using various pairs of spectral colors, whose mixtures fell along the lines indicated in the chromaticity diagram of Fig. 2. (Wh in Fig. 2 shows the chromaticity of the reference white.)

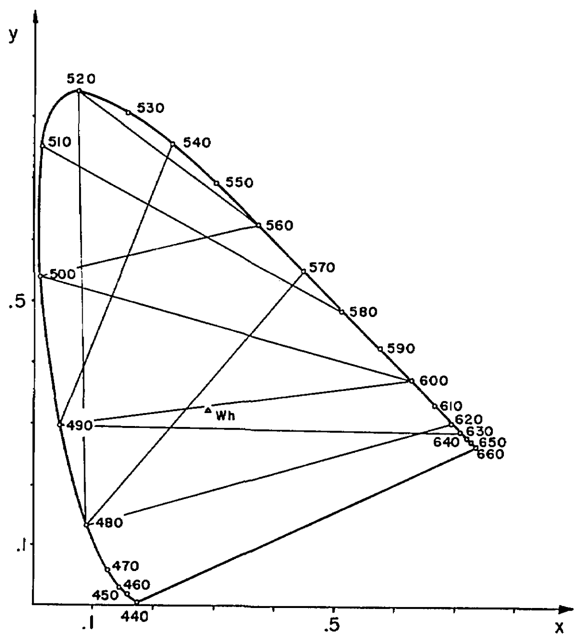

FIG. 2. CIE chromaticity diagram showing mixture lines for the various spectral colors used in the experiment, and the location of the white reference stimulus in the chromaticity plane.

Results

The resulting plots of α vs β for some of the combinations tested are depicted in Fig. 3, which shows conclusively that the results using the MDB criterion are additive because the points obtained are scattered irregularly about the ideal line. The amount of variability is of the same order as in the more limited examples previously reported by Boynton and Kaiser.[7] This was true for all combinations tested, and the results did not differ according to whether λ_1 or λ_2 was adjusted.

SPECTRAL LUMINOUS EFFICIENCY BY FOUR METHODS

Procedure

These experiments were carried out with four observers having normal color vision. Refractive errors were corrected with the use of appropriate lenses. As criteria, we used the MDB, two varieties of direct heterochromatic photometry (direct comparison and step by step) and also the minimum-flicker method (flicker photometry). The observers were presented with spectral stimuli from 400 to 670 nm at intervals of 10 nm. They were instructed to adjust the radiance of the reference white, to satisfy the criterion under consideration, with respect to the test wavelength. The spectral lights were presented at the highest radiances obtainable; maximum relative radiance is shown as a function of wavelength in Fig. 4. The results were then normalized to an equal-energy spectrum for the purpose of plotting relative luminous-efficiency functions.

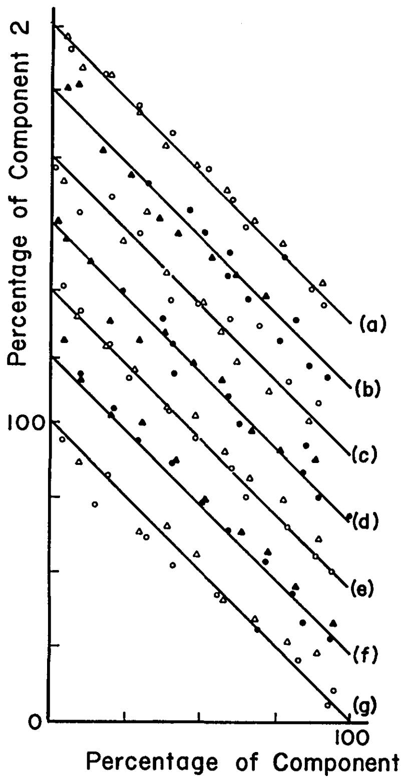

FIG. 3. Additivity of binary mixtures, using the MDB criterion, for seven pairs of spectral stimuli. Open circles: λ_1 was adjusted; open triangles: λ_2 was adjusted. (a) $\lambda_1=480$ nm, $\lambda_2=520$ nm; (b) $\lambda_1=480$ nm, $\lambda_2=620$ nm; (c) $\lambda_1=500$ nm, $\lambda_2=600$ nm; (d) $\lambda_1=510$ nm, $\lambda_2=580$ nm; (e) $\lambda_1=480$ nm, $\lambda_2=570$ nm; (f) $\lambda_1=500$ nm, $\lambda_2=540$ nm; (g) $\lambda_1=490$ nm, $\lambda_2=540$ nm.

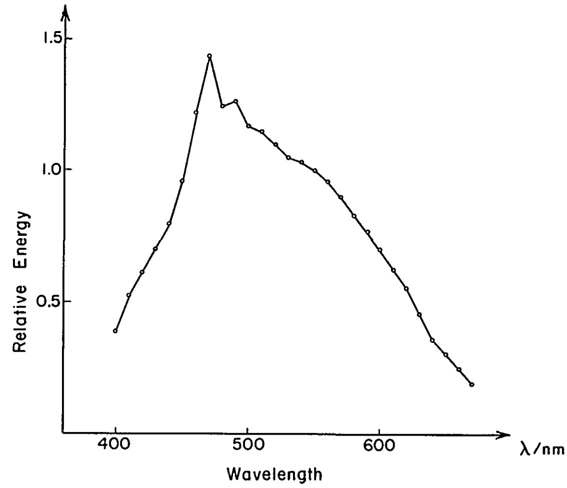

FIG. 4. Maximum obtainable relative radiance, plotted as a function of wavelength.

In the course of each experimental session, the spectrum was measured once. At each wavelength, the observer made five adjustments per session. The experimental sessions were repeated three or four times for each observer, for each criterion. Because of the limited radiances available, it was not possible to maintain high photopic levels at the extreme wavelengths of the spectrum. The maximum available retinal illuminances for various wavelengths are given in Table I. Although the illuminances at the spectral extremes are certainly mesopic, and not photopic as defined by the CIE, the use of very small fields, directly fixated, ensures that the experiments are not complicated by the intrusion of scotopic (rod) vision.

Results

The resulting spectral-luminous-efficiency curves are shown in Figs. 5–8. Each figure shows results for one of the four criteria used. In each case, the points are the

TABLE I. Retinal illuminance in photopic trolands corresponding to maximum available radiance at selected wavelengths.

Wavelength (nm)	Retinal illuminance (td)
445	10
485	100
560	574
625	100
660	10
690	1

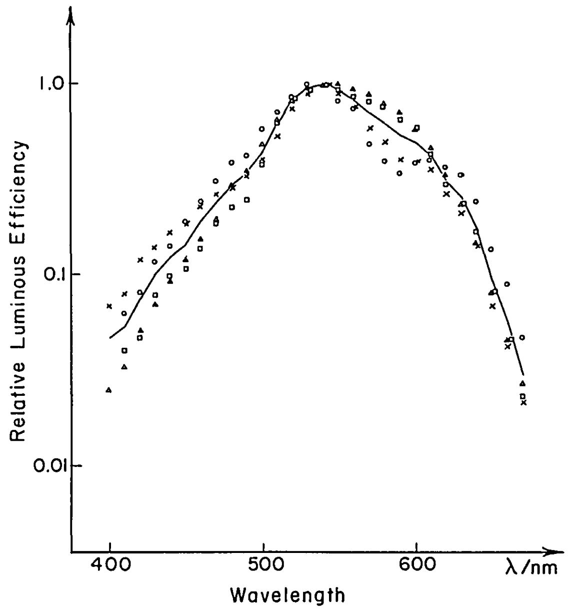

FIG. 5. Relative spectral-luminous-efficiency functions for four observers: open triangle—HGW; open square—RD; open circle—MS; cross—TSG. The solid function represents the mean data of the four observers. Criterion: equal brightness (step by step).

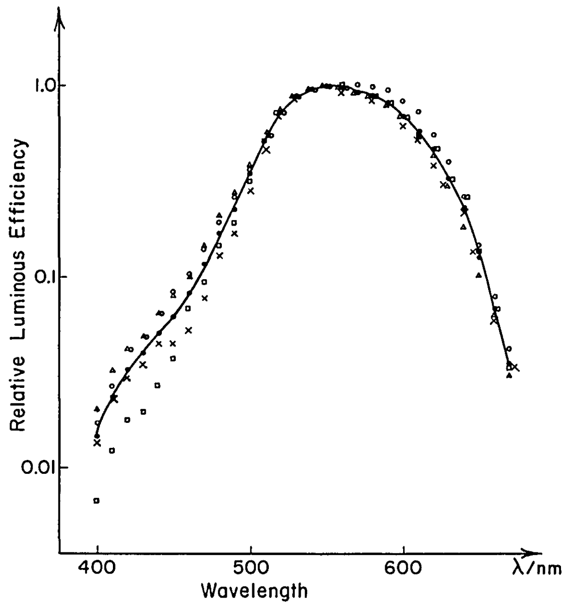

FIG. 7. Same as Fig. 5, but for the MDB criterion.

averaged data of each observer, and an over-all mean curve is drawn through them.

Figure 9 shows the curves from Figs. 5 through 8, plotted together for easy comparison. Figures 10 and 11 compare the over-all averages for the minimum-border and minimum-flicker criteria with the CIE standard observer (as modified by Judd[9]). Figures 12–14 show the results given in Figs. 5–7, replotted in a normalized representation. Here the luminous-efficiency values obtained from three of the methods (MDB, direct comparison, step by step) are divided by the corresponding values, at each wavelength, obtained by flicker photometry. This permits the use of an expanded ordinate, and makes it easier to assess deviations from the minimum-flicker curve. A criterion that yielded the same relative spectral luminosity curve as the minimum-flicker method would, on this plot, be represented by a horizontal straight line. This is very nearly true for the MDB method, and would also be approximately true

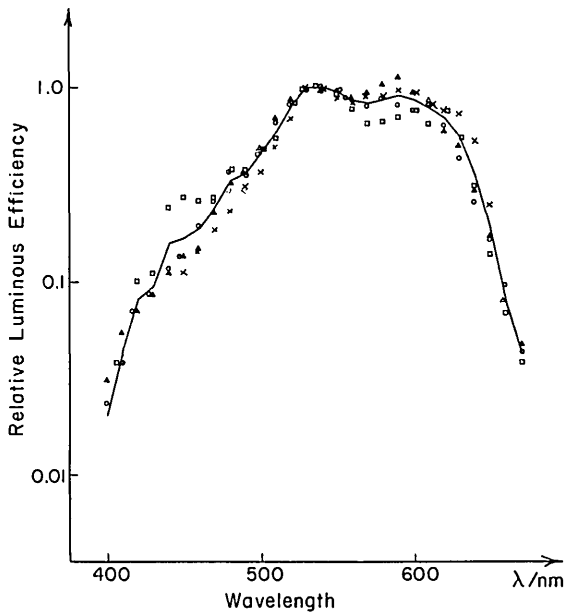

FIG. 6. Same as Fig. 5, but for the equal-brightness (direct comparison) condition.

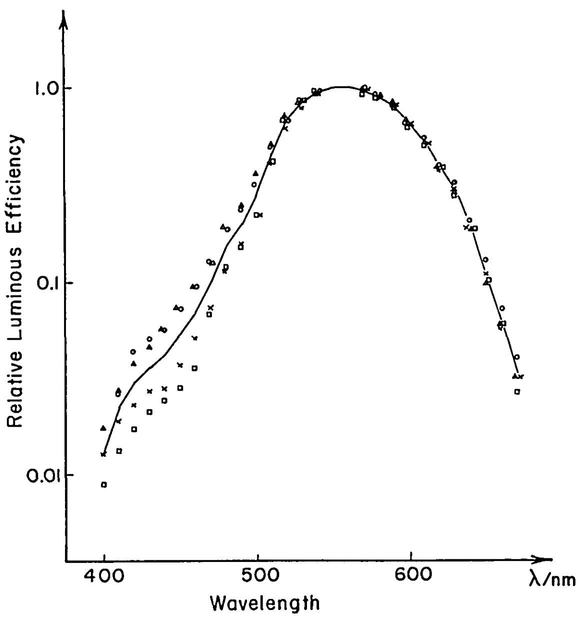

FIG. 8. Same as Fig, 5. but for the flicker criterion.

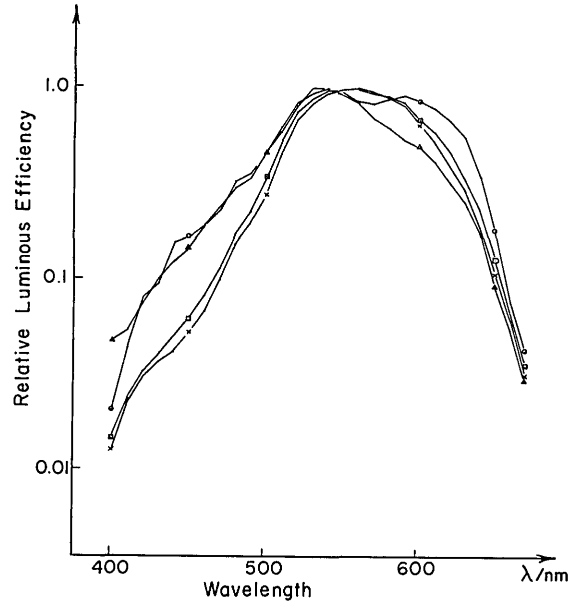

FIG. 9. Mean curves of Figs. 5–8 are replotted here upon the same graph. Open circle—equal brightness (direct comparison), open triangle—equal brightness (step by step), cross—minimum flicker, open square MDB.

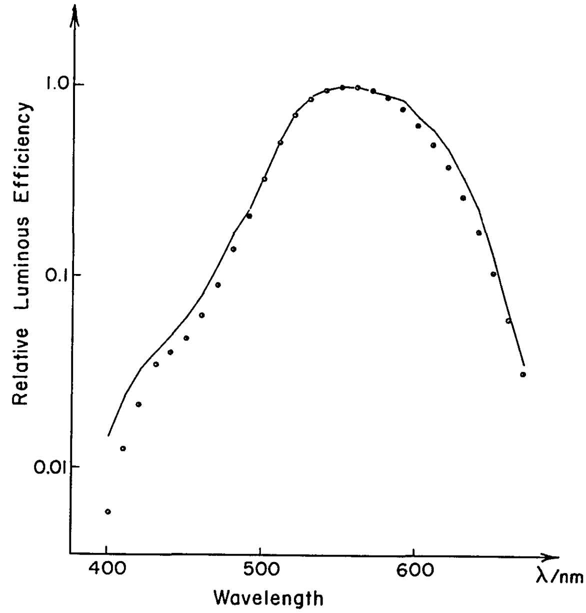

FIG. 11. Same as Fig. 10, but for the MDB criterion (solid line).

of the modified CIE data, if they were presented on this plot.

DISCUSSION

Comparison of Four Methods

The results using the minimum-flicker method agree very well with the CIE data, as modified by Judd.[9] This is not surprising since the CIE standard observer was originally based, for the most part, upon flicker-photometric measurements. The result is useful at least for confirming the adequacy of our procedures for measuring the relative radiances of the stimuli used, of our calibrations for wavelength, and for providing a basis for comparing the results of the other methods.

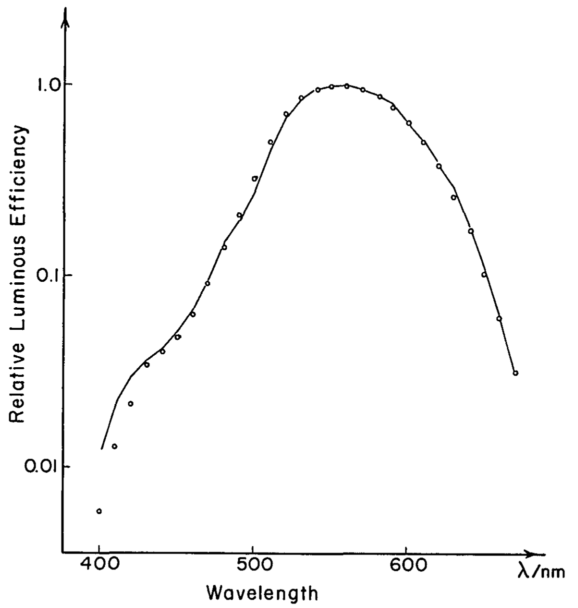

FIG. 10. Results for the minimum-flicker criterion (solid line) compared to the CIE standard observer, as modified by Judd (circles).

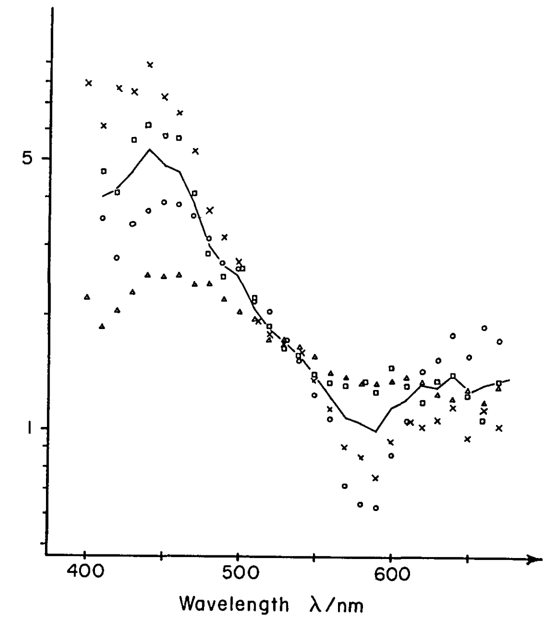

FIG. 12. Luminous efficiency for the equal-brightness (step-by-step) criterion, plotted relative to that for the flicker criterion for four subjects: open triangle—HGW; open square—RD; open circle—MS; cross—TSG. The ratio is set to 1.0 at the minimum of the function (solid curve) describing the mean values for the four observers.

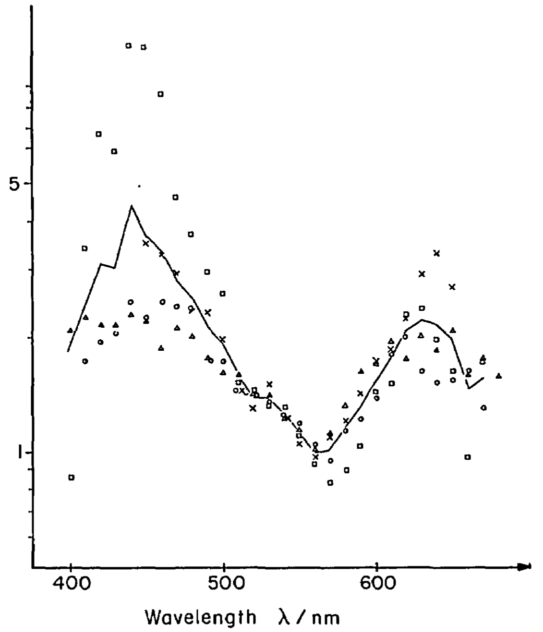

Fig. 13. Same as Fig. 12, but for the equal-brightness (direct comparison) condition.

The average results for the MDB method are obviously not different, statistically, from those obtained by flicker photometry, because the differences between the mean curves are much smaller than the differences between the results from the four subjects when either method is used. The agreement between the two methods applies to individual subjects, as well as to the overall average curves. Perhaps the only significant exception to this is a tendency for the points of subject MS to be somewhat higher in the region from 560 to 620 nm, when the MDB method is used.

It is equally clear that the results from the two methods of direct comparison agree poorly with the results from the other methods. Direct heterochromatic matches against white agree fairly well with the step-by-step method in the short-wave part of the spectrum, but at the longer wavelengths the values for step-by-step method are lower.

For relative spectral sensitivity data, comparisons of the sort just described are made difficult by the arbitrary procedure of normalizing all curves at some value, in this case at 550 nm. But the form of the difference curves shown in Figs. 11–13 is not altered thereby.

Figure 13, which shows the equal-brightness (direct-comparison) data relative to those obtained with flicker photometry, is approximately the same as would be obtained if the data were normalized against MDB method instead (because the agreement between flicker photometry and MDB is so close). The least-saturated wavelength in the spectrum is 570 nm. In other experiments[8] we have found that, when this wavelength (570 nm) is compared with white, the minimum border occurs very near to the setting that also yields equal brightness. In Fig. 13 the minimum value is very close to 570 nm. The difference between any other part of the curve and this minimum therefore very nearly represents an index of the extra brightness produced by stimuli of other wavelengths when compared with white. Specifically, it shows the ratio by which the radiance of the white reference stimulus must be increased, above the MDB value, in order to make it as bright as the spectral stimulus with which it is being compared. There are two maxima, one at about 430 nm and the other at about 630 nm.

Theory

In agreement with Boynton and Kaiser,[7] we find the variability of settings with the equal-brightness method to be significantly greater than with the MDB method. There is no difference in this respect between the MDB method and flicker photometry. These findings are illustrated in Fig. 15. The results of this experiment are also in general agreement with some recent results reported by Kaiser.[10] The results are generally what would be expected from the theory of Boynton and Kaiser[6] (see also Kaiser *et al.*[8]) except for the downturns at the spectral extremes, which may be related in some way to the lower luminances that were used in this experiment.

The Kaiser–Herzberg–Boynton theory[8] is that two stimuli produce a MDB with respect to one another because the numbers of W brain elements activated (mediating the sensation of white) are equal in response to the two parts of the field. W elements are assumed to give rise to sensations of brightness, but so are C (chromatic) elements: The excess of active chromatic elements in the more saturated field is thought to be responsible for the extra brightness that these fields elicit. Additivity failures result when opponent-color cancellation in the retina turns off chromatic brain elements that previously had been active in response to the components of the stimulus, presented individually.

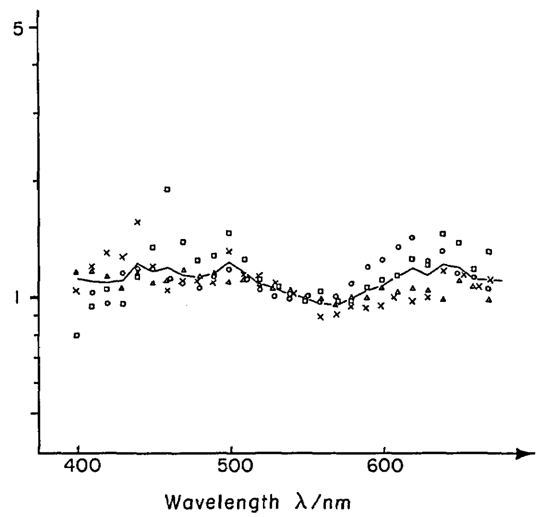

Fig. 14. Same as Fig. 12, but for the MDB criterion.

Brightness Prediction

A number of studies seem to point the way toward a solution of the problem of brightness prediction. These include the papers of Dresler,[11] Sanders and Wyszecki,[12] Kaiser,[10] Guth *et al.*,[13] and the present work. The solution of the problem seems to be to divide it into three parts. The suggestion to be made here is very similar to one advocated by Guth. In effect, Guth uses the additivity failures observed in direct brightness matching to infer the amount and type of chromatic contribution being made by each component. Guth then distinguishes between "vector luminance," which is not additive, and ordinary luminance, which is. The former is essentially what brightness matching yields, whereas the MDB or flicker methods provide the latter.

The first step is to predict the responsiveness of the achromatic system by the application of Eq. (1) or, for a particular observer, to measure it by the MDB method (or by flicker photometry, which yields very nearly equivalent results). The second step is to obtain some measure of saturation for each of the two stimuli. In our previous work, we have shown two ways to do this. One is to assess the distinctness of the minimally distinct border by a border-matching technique.[8] The other[14] is to add white light to a chromatic field, while keeping the total amount of light constant, until the MDB is driven down to some low criterion of distinctness. This second technique begins by having the experimenter add a certain amount of white light to the chromatic field, after which the subject adjusts the luminance of the mixture field, without altering the ratio of the luminances of the components. This mixture field, juxtaposed with white, yields a minimally distinct border whose distinctness is then evaluated by the first technique. By repeating the experiment with various amounts of white light, a condition can be found for which the amount of added white reduces the border distinctness to criterion. The more saturated the chromatic field, the more white light must be added to yield the criterion border distinctness.

The third step is to develop a predictive scheme whereby the extra brightness to be expected from the more-saturated stimulus can be computed, preferably to predict the factor by which its radiance must be reduced to make it appear as bright as the less-saturated stimulus with which it is juxtaposed.

For the evaluation of spectral sensitivity according to a procedure that yields additive results, one that presumably evaluates only the achromatic activity of the visual system, it seems according to the results of our study that either flicker photometry or MDB could be used equally well, since the methods yield very similar results. But for obtaining a measure of saturation, the MDB method is superior, because there are at least two ways to use the MDB method to obtain an index of saturation, whereas flicker photometry does not provide such information at all.

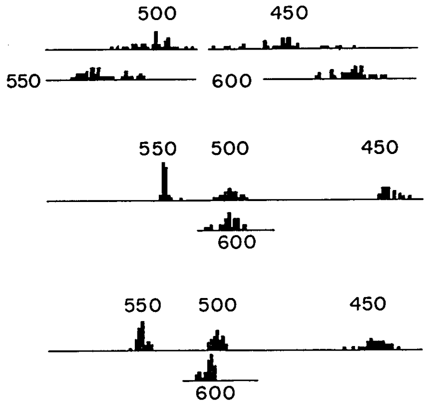

FIG. 15. Frequency distribution of neutral density wedge settings. Four wavelengths are selected, and the spread of the results for equal brightness matches (direct comparison), MDB, and flicker are shown in the top, middle, and bottom sections of the figure, respectively. The relative horizontal displacement of the various distributions is without significance. Class intervals are 0.02 density units.

ACKNOWLEDGMENTS

This work was supported by Grant Nos. NB-00624 and EY-00187 from the National Institute of Neurological Diseases and Blindness and the National Eye Institute. Dr. Wagner was supported by a postdoctoral fellowship (DAAD, Ref. 4-nato-3-kü-ht).

REFERENCES

[1] See Y. LeGrand, *Light, Colour and Vision, Second ed.*, translated by R. W. G. Hunt, J. W. T. Walsh, and F. R. W. Hunt (Chapman and Hall, London, 1968), Ch. 6; also J. W. T. Walsh, *Photometry, 3d ed.* (Dover, New York, 1965), for detailed discussions of the problems raised here.

[2] G. Heath, Science **128**, 775 (1958).

[3] J. L. Brown, L. Phares, and D. E. Fletcher, J. Opt. Soc. Am. **50**, 950 (1960).

[4] H. E. Ives, Phil. Mag. **24**, 149 (1912).

[5] J. Guild, in *Discussion on Vision* (Physical Society, London, 1932); in *Sources of Color Science*, edited by D. L. MacAdam (MIT Press, Cambridge, Mass., 1971), p. 213.

[6] The bracketed portions of this quotation have been supplied by us.

[7] R. M. Boynton and P. K. Kaiser, Science **161**, 366 (1968).

[8] P. K. Kaiser, P. A. Herzberg, and R. M. Boynton, Vision Res. **11**, 953 (1971).

[9] D. B. Judd, in *CIE* (Commission Internationale de l'Eclairage) *Proceedings, 12th Session Stockholm, 1951* (Bureau Central CIE, 57 rue de Pépinère, Paris, 1952), Vol. 1, Pt. 7, p. 11.

[10] P. K. Kaiser, J. Opt. Soc. Am. **61**, 966 (1971).

[11] A. Dresler, Trans. Illum. Eng. Soc. (London) **18**, 141 (1953).

[12] C. L. Sanders and G. Wyszecki, in *CIE* (Commission Internationale de l'Eclairage) *Proceedings, 15th Session Vienna, 1963* (Bureau Central CIE, 57 rue de Pépinère, Paris, 1964), Vol. B, p. 221.

[13] S. L. Guth, N. J. Donley, and R. T. Marrocco, Vision Res. **9**, 537 (1969).

[14] R. M. Boynton and T. S. Greenspon, Vision Res. **12**, 495 (1972).