

Theory of wavelength discrimination in tritanopia

Kenneth Knoblauch

Vision Research Laboratory, The Lighthouse Inc., 800 Second Avenue, New York, New York 10017

Received July 13, 1992; accepted September 14, 1992; revised manuscript received September 28, 1992

Many theories of color discrimination predict a discontinuity in the wavelength-discrimination function of a tritanope at the point in the spectrum at which the rate of change of the visual signal constrained to an equilibrium plane passes through zero (near 460 nm). The predicted discontinuity follows from the use of a first-order approximation for which the reciprocal of the slope of the response function that generates the visual signal is proportional to the discrimination limen. In view of the good discrimination shown by such observers elsewhere in the spectrum, however, such a singularity is impossible. I show that the inclusion of the higher-order terms produces a finite value in the 460-nm region that falls in the range of values from the literature that have been obtained experimentally.

INTRODUCTION

Nearly 50 years ago the first reports appeared that documented the effects of the loss of the short-wave (*S*) cone response on visual function.¹⁻³ One striking finding in these early reports was the sharp deterioration in wavelength discrimination under such conditions in the vicinity of 460 nm. The just-noticeable differences (jnd's) for Wright's tritanopic observers, for example, were ~50 nm.³ Wright referred to this drastic loss in sensitivity as a "discontinuity in the blue-green wavelengths where wavelength discrimination virtually disappears" Subsequent reports by numerous investigators⁴⁻¹⁰ on additional tritanopic observers or under conditions that minimized the contribution of the *S* cones confirm Wright's findings, although discrimination losses of the magnitude reported by Wright were not typically found (Table 1). Deterioration of discrimination in the 460-nm region with little change throughout the rest of the spectrum is now taken as indication of the failure of the *S* cones to contribute to specific types of discrimination.⁹⁻¹² Mollon and Estévez¹¹ termed the peak in loss of discrimination that occurs in the 460-nm region as well under certain conditions in normal vision the short-wave pessimum.

In the literature since Wright, there is a flaw in the interpretation that is commonly given to the loss of discrimination. The quotation from Wright refers to the loss of discrimination as a discontinuity, and indeed Wright draws a curve through the data with a singularity near 460 nm. The implication is that wavelength discrimination becomes infinitely poor here. But such behavior is impossible, as just outside both sides of this region discrimination is nearly as acute as that of a normal observer. Wavelength discrimination can be poor, but not infinitely so at a single point within the spectrum. Another sentiment sometimes expressed, that wavelength discrimination is indeterminate in the 460-nm region, indicates no more than the tautology that discrimination is indeterminate within a jnd. The implication that the wavelength-discrimination curve for a tritanope is discontinuous (or unmeasurable) in one part of the spectrum is often repeated in more modern contexts,^{4,13} and theoretic-

cal curves that imply such a situation can be found in numerous sources.^{8,12-14}

In the present Communication it is shown that the theory typically used to model wavelength discrimination in the absence of *S* cone signals involves a simplifying assumption that is invalid in the vicinity of the short-wave pessimum and that is responsible for the singularity portrayed by many investigators. An alternative formulation will be introduced that gives finite values for discrimination in the region in question that are in accord with the bulk of the empirical data. This alternative formulation, when interpreted within the framework of the wavelength-discrimination procedure itself, suggests why tritan observers experience difficulty in performing the task in the 460-nm spectral region.

THEORY

Discrimination Models

In a typical wavelength-discrimination experiment a bipartite field is presented to the observer. On one half of the field a standard wavelength λ is presented, and on the other half a comparison wavelength $\lambda + \Delta\lambda$ is presented. The two half-fields are maintained at equal luminance, and the difference in wavelength between them, $\Delta\lambda$, is increased until the observer can discriminate the change according to some criterion. It is assumed that the difference between the two half-fields is just discriminable when

$$f(\lambda, \Delta\lambda) = |\Gamma(\lambda) - \Gamma(\lambda + \Delta\lambda)| = c, \quad (1)$$

where c is a criterion value and Γ , the visual signal, is a real, single-valued function of wavelength that is bounded (i.e., finite) over the visible spectrum.

If Γ is differentiable with a bounded derivative, then Eq. (1) can be rewritten in terms of the definite integral,

$$|\Gamma(\lambda) - \Gamma(\lambda + \Delta\lambda)| = \left| \int_{\lambda}^{\lambda + \Delta\lambda} \Gamma'(\lambda) d\lambda \right| = c. \quad (2)$$

The wavelength-discrimination function is a level curve of the function $f(\lambda, \Delta\lambda)$ defined by Eqs. (1) and (2). One

Table 1. Summary of Characteristics of Discrimination at Short-Wavelength Pessimum Found under Tritan Conditions

Ref.	Location (nm)	$\Delta\lambda$ (nm)	Condition	Method
1	≈460	≈50	Small field	Estimated midpoint of range
2	450-480	25	Tritanopia	Average
3	≈465	≈50	Tritanopia	Estimated midpoint of range
4	430-450	20-30	Tritanopia	Steps toward long wave
5	480	≈10	Tritanopia	Slope of color naming
6	440-480	—	Tritanopia	—
7	450-460	10-20	Small field	—
8	455-480	10	Tritanopia	Average
9	460	25-35	Transient tritanopia, Just-noticeable difference and just-noticeable border	Average
10	460-475	35-45	Small field, just-noticeable inhomogeneity	Average

could solve Eq. (2) for $\Delta\lambda$, but it is conceptually equivalent and usually simpler to interpolate directly on the curve $\Gamma(\lambda)$ to find the values of $\Delta\lambda$ that yield a criterion difference.

In common practice, however, one approximates Eq. (1) by assuming that $\Delta\lambda$ is small, in which case, based on the definition of the derivative, Eq. (1) simplifies to

$$\Delta\lambda|\Gamma'(\lambda)| = c, \tag{3}$$

which is easily solved for $\Delta\lambda$,

$$\Delta\lambda = \frac{c}{|\Gamma'(\lambda)|}. \tag{4}$$

If $\Gamma' = 0$, $\Delta\lambda$ becomes infinite, which would violate the assumption that $\Delta\lambda$ is small. Equivalently, Eq. (4) can be derived from the Taylor series expansion of Γ about λ :

$$\Gamma(\lambda + \Delta\lambda) = \Gamma(\lambda) + \Gamma'(\lambda)\Delta\lambda + \Gamma''(\lambda)\Delta\lambda^2/2! + \dots \tag{5}$$

Over most of the spectrum the terms of higher order than the first derivative are likely to be small, and when these terms are ignored, Eq. (5) can be cast in the form of Eq. (4). When the first derivative is near zero, however, the quadratic and higher terms control the difference, $\Gamma(\lambda) - \Gamma(\lambda + \Delta\lambda)$, which will produce results that are inconsistent with the assumption that this difference is well approximated by $\Delta\lambda\Gamma'(\lambda)$. On the other hand, solving the integral equation or interpolating on the response function is always correct, since neither procedure involves an approximation.

Color Vision Model

It is assumed that in tritanopia (or under conditions mimicking tritanopia) there are no functioning *S* cones. In addition, under the conditions of the wavelength-discrimination experiment, variation in the stimulus produces no change in the luminance. If I_λ is the energy of a monochromatic light of wavelength λ , then $I_\lambda(L + M) = \text{constant}$, where *L* and *M* are, respectively, the spectral sensitivities of the long-wavelength- and midwavelength-sensitive cone systems weighted to add to the luminance spectral sensitivity and with their wavelength dependence implicit. For simplicity the constant is assumed to be unity. Then the energy, I_λ , necessary to maintain a con-

stant luminance across wavelengths is $(L + M)^{-1}$. Given the above considerations, the discrimination function is assumed to depend only on the chromatic pathway mediated by antagonistic interactions between *L* and *M* cones. This signal is represented as

$$\Gamma(\lambda) = (L - 2M)/(L + M), \tag{6}$$

where the explicit use of Γ here clarifies the connection with the discrimination models presented above. The factor of 2 applied to the *M*-cone spectral sensitivity in the numerator represents the approximate weighting necessary to produce a spectral zero crossing near 570 nm for the chromatic system when the *L* and *M* signals are preweighted to sum to the luminance spectral sensitivity. The constraint that all lights be of equal luminance means that the *M*-cone signal does not change independently of the *L*-cone signal: $M = (1 - I_\lambda L)/I_\lambda$. Substitution of this expression for *M* and rearrangement of terms yields the following equation, which indicates that the signal can be represented as a function of the *L*-cone signal restricted to the equal-luminance plane:

$$\Gamma(\lambda) = -2 + 3I_\lambda L. \tag{7}$$

The wavelength-discrimination function generated from Eqs. (4) and (7) displays a singularity at ~465 nm and a minimum near 600 nm. Psychophysical data seem to show a minimum nearer 580 nm. In recent treatments^{13,15,16} an additional term is introduced that incorporates into the equations the fact that opponent mechanisms are less sensitive away from their equilibrium point. For example, following Boynton,¹³ the right-hand side of Eq. (4) could be multiplied by $1 + |L - 2M|$ (termed a *J* factor by Boynton), which has the effect of increasing the jnd everywhere except at the equilibrium wavelength. Such a term can be interpreted as resulting from response compression or adaptation that no doubt occurs in a standard wavelength-discrimination paradigm when the fields are presented continuously.

Wavelength discrimination based on a signal using Eqs. (4) and (7) and the *J* factor is shown in Fig. 1 as the solid curve. The points through which the smooth curve is interpolated are from Table 8.1, column F, of Ref. 13. It is evident that, although the cone antagonistic term does

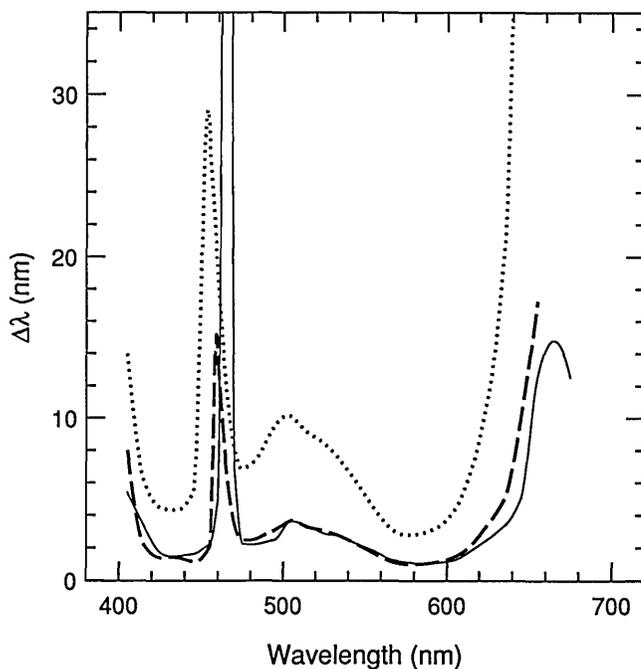


Fig. 1. Theoretical curves of the just-discriminable wavelength difference $\Delta\lambda$ as a function of wavelength for a tritanope. The solid curve is based on the assumption that discrimination is proportional to the slope of the visual response curve. The dashed and dotted curves were calculated by interpolation of criterion values of 0.01 and 0.03, respectively, along the visual response curve.

differentially weight the wavelength-discrimination curve so that it is in qualitative accord with human data over most of the spectrum, it does not eliminate the singularity near 465 nm. This wavelength-discrimination curve is based on the assumption that the jnd is proportional to the reciprocal of the slope (estimated by a finite difference). As demonstrated above [Eq. (4)], the condition on which such an assumption is based becomes invalid as the slope approaches zero.

RESULTS

The dashed and dotted curves in Fig. 1 were generated for two criterion levels by interpolating the wavelength difference that yielded a criterion value along the curve, $\Gamma(\lambda)$, generated from Eq. (7), preweighted by the reciprocal of the J factor described above. The Smith-Pokorny fundamentals were used for the L - and M -cone sensitivities (tabulated in Ref. 13), and the weighted samples of Γ were interpolated with a cubic spline.¹⁷ Over most of the spectrum the calculations were performed for standard wavelengths at 10-nm intervals, stepping toward only the long-wavelength end of the spectrum, rather than averaging thresholds in the long- and short-wave directions. In the vicinity of 460 nm, values were calculated at 1-nm intervals to locate the pessimum more accurately.

The dashed curve represents a criterion value of 0.01. Over most of the spectrum the dashed curve matches the solid curve well, which indicates in these regions the accuracy of the derivative approximation given by Eq. (4). However, for large values of $\Delta\lambda$, near 460 and above 620 nm, the curves deviate from each other significantly.

A pessimum occurs for the dashed curve at 459 nm. The exact peak depends on the criterion used, however, with a larger criterion value giving a pessimum at shorter wavelengths. The dotted curve in Fig. 1, which shows such an effect, was generated with a criterion three times as large. (If the thresholds in both directions are averaged with either criterion, the pessimum moves to 460 nm.) Increasing the criterion in this fashion increases the pessimum by a factor of 2 and results in poorer performance throughout the spectrum. The dashed and dotted curves span most of the range of values found in the literature since Wright for tritanopic wavelength-discrimination (Table 1). Although these results show the effect that a criterion can have on the curve, stimulus and procedural differences also may play an important role in determining the size of the jnd measured.

DISCUSSION

The benefits of using Eq. (2) instead of an approximation are that the derivative does not appear in the denominator and infinities are avoided. An approach such as the one used here was followed by Hurvich and Jameson¹⁸ to derive wavelength-discrimination curves of dichromats. Cavanaugh and Estévez⁷ also obtained finite values for discrimination by calculating the change in the ratio of π -mechanism sensitivities for a criterion response. Many line element models,¹⁹⁻²² on the other hand, incorrectly predict a singularity in the tritanopic wavelength-discrimination curve because the finite difference $\Delta\lambda$ is treated as if it were a differential $d\lambda$. Other models^{23,24} have formally treated the discrimination limen in this region so that it would be finite. Nevertheless, the predicted curves that have been published were drawn with a singularity near 460 nm.

The theoretical procedure described here provides a more natural interpretation of the tritanopic pessimum. Instead of attributing the discrimination loss to the fact that the slope of the spectral signal becomes zero at a specific point in the spectrum, poor performance in the 460-nm spectral region can be related to the existence of metameric pairs of monochromatic lights straddling 460 nm.²⁵ For example, if the standard wavelength is less than 460 nm and within a distance from that wavelength that produces less than a criterion response, then, as the observer turns the monochromator dial, the difference in physiological response between the two half-fields becomes greater up to 460 nm but not so great as to allow the two sides of the field to be discriminable. As the observer continues to increase the wavelength separation, the physiological difference between the two fields decreases up to the point at which the metameric wavelength is reached. Then, for the two halves of the field to be discriminable from each other, the observer must continue to increase the wavelength difference until a criterion value is reached.

By this explanation the wavelength limen, as measured starting slightly above or below 460 nm and moving toward 460 nm, should be greater than the limen starting at 460 nm and moving in either direction. This explanation may account for differences in the position and the size of the pessimum among studies that used different methods for measuring the jnd and assigning its value to a particu-

lar wavelength. For example, Krill *et al.*⁴ measured steps only in the long-wave direction and found a pessimum in the range 430–450 nm, which is somewhat shorter than that found in other studies that averaged steps toward the short- and the long-wavelength regions or estimated the midpoint of the range from steps in both directions.

Explaining the deterioration in discrimination in a similar fashion, Wilmer and Wright¹ argued that the non-monotonic variation in the *quality* of the light in the 460-nm spectral region results in the observations' having little meaning. Such changes in quality are by definition subliminal, however, and so it is difficult to see how they would invalidate the measurement of the size of the wavelength difference that the individual can discriminate in this part of the spectrum. The difficulty is that experienced by the observer performing the task in this region. In the vicinity of 460 nm some confusion may be exhibited because the change in appearance is similar for a change in wavelength in both directions. Mollon *et al.*¹² describe normal observers who experience exactly this difficulty at equiluminance when the pathway through which the *S* cones send their signals is saturated and contributes little to discrimination.

In summary, the commonly used approximation that *jnd*'s are proportional to the reciprocal of the slope of the underlying response function depends on the assumption that such *jnd*'s are small relative to the response. Under such conditions higher-order terms of the series expansion of the response function can be safely ignored. Such an approximation is likely to be valid for normal trichromats who have fine wavelength discrimination over a large part of the visible spectrum. When sensitivity deteriorates, however, as it does at the short-wavelength pessimum in tritanopia, one must take such higher-order terms into account in order to represent discrimination behavior sensibly and accurately.

ACKNOWLEDGMENT

The author thanks Aries Arditi, Rhea T. Eskew, Laurence T. Maloney, and Joel Pokorny for their helpful suggestions and critical comments on an earlier draft of this note. This work was supported by National Institutes of Health grant EY07747.

REFERENCES AND NOTES

1. E. N. Wilmer and W. D. Wright, "Colour sensitivity of the fovea centralis," *Nature (London)* **156**, 119–121 (1945).
2. F. P. Fischer, M. A. Bouman, and J. ten Doesschate, "A case of tritanopy," *Doc. Ophthalmol.* **5/6**, 73–87 (1951).
3. W. D. Wright, "The characteristics of tritanopia," *J. Opt. Soc. Am.* **42**, 509–521 (1952).
4. A. E. Krill, V. C. Smith, and J. Pokorny, "Further studies supporting the identity of congenital tritanopia and hereditary dominant optic atrophy," *Invest. Ophthalmol.* **10**, 457–465 (1971).
5. D. Smith, "Color naming and hue discrimination in congenital tritanopia and tritanomaly," *Vision Res.* **13**, 209–218 (1973).
6. J. Voke-Fletcher and R. J. Fletcher, "A case of tritanopia," *Mod. Probl. Ophthalmol.* **19**, 229–231 (1978).
7. C. R. Cavonius and O. Estévez, " π -mechanisms and cone fundamentals," in *Visual Psychophysics and Physiology*, J. C. Armington, J. Krauskopf, and B. R. Wooten, eds. (Academic, New York, 1978), pp. 221–231.
8. M. Alpern, K. Kitahara, and D. H. Krantz, "Classical tritanopia," *J. Physiol.* **335**, 655–681 (1983).
9. P. K. Kaiser and R. M. Boynton, "Role of the blue mechanism in wavelength discrimination," *Vision Res.* **25**, 523–529 (1985).
10. P. K. Kaiser and M. Ayama, "Just noticeable inhomogeneity criterion for determining wavelength discrimination function," *Vision Res.* **25**, 1327–1330 (1985).
11. J. D. Mollon and O. Estévez, "Tyndall's paradox of hue discrimination," *J. Opt. Soc. Am. A* **5**, 151–159 (1988).
12. J. D. Mollon, O. Estévez, and C. R. Cavonius, "The two subsystems of colour vision and their rôles in wavelength discrimination," in *Vision: Coding and Efficiency*, C. Blakemore, ed. (Cambridge U. Press, Cambridge, 1990), pp. 119–131.
13. R. M. Boynton, *Human Color Vision* (Holt, Rinehart & Winston, New York, 1979).
14. J. Pokorny, V. C. Smith, G. Verriest, and A. J. L. G. Pinckers, *Congenital and Acquired Color Vision Defects* (Grune & Stratton, New York, 1979).
15. R. M. Boynton and N. Kambe, "Chromatic difference steps of moderate size measured along theoretically critical axes," *Color Res. Appl.* **5**, 13–23 (1980).
16. A. L. Nagy, R. T. Eskew, and R. M. Boynton, "Analysis of color-matching ellipses in cone-excitation space," *J. Opt. Soc. Am. A* **4**, 756–768 (1987).
17. The differentiable approximations of *L* and *M* cone sensitivities published by R. M. Boynton and J. J. Wisowaty ["Equations for chromatic discrimination models," *J. Opt. Soc. Am.* **70**, 1471–1476 (1980)] could have been used here, although their use would be unlikely to have affected the results. Contrary to the authors' assertion, the cubic spline curve is continuous [W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in Pascal, The Art of Scientific Computing* (Cambridge U. Press, Cambridge, 1989), Sec. 3.3] and should suffice for the purposes here.
18. L. M. Hurvich and D. Jameson, "Some quantitative aspects of an opponent-colors theory. II. Brightness, saturation, and hue in normal and dichromatic vision," *J. Opt. Soc. Am.* **45**, 602–616 (1955).
19. W. S. Stiles, "A modified Helmholtz line-element in brightness-colour space," *Proc. Phys. Soc.* **58**, 41–65 (1946).
20. M. A. Bouman and P. L. Walraven, "Quantum theory of discrimination of dichromates," *Vision Res.* **2**, 177–187 (1962).
21. J. J. Vos and P. L. Walraven, "An analytical description of the line element in the zone-fluctuation model of colour vision—I. Basic concepts," *Vision Res.* **12**, 1327–1344 (1972).
22. J. J. Vos and P. L. Walraven, "An analytical description of the line element in the zone-fluctuation model of colour vision—II. The derivation of the line element," *Vision Res.* **12**, 1345–1365 (1972).
23. D. B. Judd and G. T. Yonemura, "CIE 1960 UCS diagram and the Müller theory of color vision," *J. Res. Nat. Bur. Stand. Sect. A* **74**, 23–30 (1970).
24. S. L. Guth, R. Massof, and T. Benzschawel, "Vector model for normal and dichromatic vision," *J. Opt. Soc. Am.* **70**, 187–211 (1980).
25. Given that the slope is zero, such metameric pairs need not exist; if the second derivative vanishes also, then the point at which the slope becomes zero is an inflection rather than an extremum, and the spectral curve would be monotonic.