

COLOR METRICS

INSTITUUT VOOR
ZINTUIGFYSIOLOGIE TNO
Soesterberg
Kampweg 5 - Postbus 23
Tel. 0468 - 1444

Helmholtz Memorial Symposium on Color Metrics

Organizing Committee

P. L. Walraven, chairman
L. F. C. Friele
D. van Norren
J. L. Ouweltjes
J. J. Vos

Copyright 1972 reserved AIC/Holland, Soesterberg.

All rights reserved. No part of this book may be reproduced by any means, without written permission from the publisher.

COLOR METRICS

Proceedings of the Helmholtz Memorial Symposium on
Color Metrics, held at Driebergen, The Netherlands,
September 1-3, 1971, sponsored by the International
Color Association AIC and the Nederlandse Vereniging
voor Kleurenstudie (AIC/Holland)

edited by

J. J. VOS, L. F. C. FRIELE
and
P. L. WALRAVEN

Published by AIC/Holland, c/o Institute for Perception TNO

Soesterberg — 1972

INSTITUUT VOOR
ZINTUIGFYSIOLOGIE TNO

Soesterberg

Kampweg 5 — Postbus 23

Tel. 03465 - 1444



Hermann von Helmholtz 1821 — 1894
founder of Color-metrics
(from a lithography by Lenbach, 1894)

PREFACE

The history of this symposium dates back to the 1969 meeting of the International Color Association AIC in Stockholm. There, the Dutch National Committee proposed, and got approval, to hold a limited attendance symposium in 1971 on the topic of "small color differences". This idea agreed with wishes expressed earlier in AIC circles to hold small specialized symposia in between the large quadrennial congresses.

The Organizing Committee made two small but significant changes in the symposium title. One was to shift from "small color differences" to "color metrics" thus indicating a shift in emphasis from purely experimental to one also including more theoretical viewpoints. The second change was made when it was realized that the symposium would start — admittedly by sheer coincidence — the day after the 150th anniversary of the birth of Hermann von Helmholtz, undisputed founder of color metrics.

The "Helmholtz Memorial Symposium on Color Metrics", then, was meant to be a meeting between workers in the applied fields, confronted daily with problems of industrial color tolerances, and psychophysicists, primarily concerned with understanding basic mechanisms. Indeed, 64 participants, coming from all parts of the world, met and clashed as the initiators hoped they would do. These proceedings, with their almost verbatim recording of the discussions, may reflect the vivid spirit which linked the participants even after the symposium was formally closed.

The opening and final paper of these proceedings may well serve to mark the gradual shift from a basic research orientation on the opening day to a more and more applied one as time proceeded. The Organizing Committee was particularly pleased to find Dr. W. S. Stiles willing to introduce the history of the symposium theme with an invited lecture. The text of this lecture may in itself be regarded as a milestone in the development of color metrics. We are also indebted to Dr. G. Wyszecki, who allowed us to conclude these proceedings with a recently written review on the 1971 state of the art in colour difference formulas. Between these two landmarks the 25 contributions of other participants will be found. Everyone in his own way, reported on his own work, on his own struggles to find answers in a field where mutual understanding of what is meant by tolerance, acceptability, and perceptibility, is often hard to find.

Neither the symposium, nor the publication of these proceedings would have been possible without the help of many. We gratefully acknowledge the financial guarantee of the Ministry of Science and of the N. V. Philips' Gloeilampenfabrieken and the financial assistance of the Dutch Color Association NVVK.

We may also mention the amount of extra work done by the staff of the Institute for Perception TNO, which helped run the symposium and prepare the proceedings. To mention everyone would be impossible, but we may make an exception for a few, implicitly thanking all the others as well. Miss Riet de Groot and Miss Tonnie van der Werff were in charge of all secretarial duties; together with Miss Margreet Frederiks, librarian, they co-prepared major parts of the manuscripts. Mr. Jan Boogaard was in charge of all technical arrangements; the scientific staff of the Vision Branch took care of reporting the discussions; and Mr. André Huigen redrew many of the illustrations.

The Organizing Committee

PARTICIPANTS

- R. Ardoullie, Boulevard Prince de Liège, 35, Brussels, Belgium.
- C. J. Bartleson, Vice President for Research, Kollmorgen Corporation, Macbeth Research Laboratories, Box C, Newburgh, N. Y. 12550, U.S.A.
- F. W. Billmeyer, Rensselaer Polytechnic Institute, Troy, New York 12181, U.S.A.
- M. A. Bouman, Physical Laboratory, University of Utrecht, Eisenhowerlaan 4, Utrecht, The Netherlands.
- R. M. Boynton, Center for Visual Science, University of Rochester, Rochester, New York 14627, U.S.A.
- A. Brockes, Ing. Abt. AP 18 Bayer, 509 Leverkusen, Germany.
- F. J. J. Clarke, Metrology Centre, National Physical Laboratory, Teddington, Middlesex, England.
- E. Coates, University of Bradford, School of Colour Chemistry, Bradford 7, England.
- A. de la Cruz, Sociedad Espanola de Optica, Comite Espanol de Color, Serrano, 121, Madrid-6, Espana.
- F. L. Engel, Institute for Perception Research, Insulindelana 2, Eindhoven, The Netherlands.
- L. F. C. Friele, Fiber Institute TNO, Postbox 110, Delft 2207, The Netherlands.
- E. Ganz, CIBA-GEIGY Ltd., FO 3.21, CH-4002, Basel, Switzerland.
- P. Buchner, c/o SANDOZ A. G., Postfach 4002, Basel, Switzerland, CH-4002.
- S. L. Guth, Indiana University, Department of Psychology, Psychology Building, Bloomington, Indiana 47401, U.S.A.
- A. Hård, Swedish Colour Centre, Postbox 45020, 10430 Stockholm 45, Sweden.
- T. Seim, Laboratorium für Farbenmetrik, Physikalisches Institut der Universität Basel, Klingelbergstrasse 82, Switzerland.
- R. W. G. Hunt, Research Laboratories Kodak Ltd., Harrow, Middlesex, England.
- R. S. Hunter, President Hunter Laboratory, Inc. 9529 Lee Highway, Fairfax, Virginia 22030, U.S.A.
- T. Indow, Keio University, Mita, Minato-Ku, Tokyo, Japan.
- I. G. H. Ishak, The Paint Research Station, Waldegrave Road, Teddington, Middlesex, England.
- S. M. Jaeckel, The Hosiery and Allied Trades Research Association, Thorneywood. 7 Gregory Boulevard, Nottingham NG7 6LD, England.

- N. Kambe, Toshiba Research and Development Center, Tokyo Shibaura Electric Co., Ltd, 1 Komukai Toshiba-Cho, Kawasaki 210, Japan.
- P. Kowaliski, Kodak Pathé S. A., 30, Rue des Vignerons-Vincennes, France.
- R. D. Lozano, Laboratorio de Radiaciones, Department di Fisica, I.N.T.I., Libertad 1235, Buenos Aires, Argentino.
- D. MacAdam, U.S. Technical Committee on Colorimetry, C.I.E., Eastman Kodak Company, Rochester, New York 14650, U.S.A.
- D. J. McConnell, Bowaters Scientific Service Division, Northfleet, Kent, England.
- R. MacDonald, J. & P. Coats, Ltd., 155 St. Vincent Street, Glasgow C.2, P. O. Box 34, Scotland.
- K. McLaren, Imperial Chemical Industries Ltd., Dyestuffs Division, P. O. Box 42, Hexagon House, Blackley Manchester M9 3DA, England.
- F. Malkin, The British Ceramic Research Association, Queens Road, Penkhull Stoke-on-Trent, ST4 7 LQ, England.
- V. S. Mihajlow, Xerox Corporation, Xerox Street, Rochester, New York 14603, U.S.A.
- Y. Nayatani, Electrotechnical Laboratory, Osaka Branch, 16 Nakoji, Amagasaki, Hyogo, Japan.
- D. Nickerson, 2039 New Hampshire N. W., Washington D. C. 20009, U.S.A.
- I. Nimeroff, National Bureau of Standards, Washington D. C. 20234, U.S.A.
- D. van Norren, Institute for Perception TNO, Kampweg 5, Soesterberg, The Netherlands.
- K. Okada, Lighting Research & Advisory Bureau, Matsushita Electric Industrial Co., Lt., Kadoma, Osaka, Japan.
- J. L. Ouweltjes, Philips' Lighting Factories, Eindhoven, The Netherlands.
- F. Parra, Muséum National d'Histoire Naturelle, Laboratoire de Physique, 43 Rue Cuvier, Paris-5, France.
- M. R. Pointer, Imperial College, Applied Optics Section, London SW1, England.
- J. R. Provost, University of Bradford, School of Colour Chemistry, Bradford 7, England.
- Ch. D. Reilly, E. I. du Pont de Nemours & Comp. Ind., Wilmington, Delaware 19898, U.S.A.
- K. Richter, Bundesanstalt für Materialprüfung, Fachgruppe Farbmeterik, 1 Berlin (West) 45, Unter den Eichen 87, Germany
- M. Richter, FNF Bundesanstalt für Materialprüfung, D-1 Berlin 45, Unter den Eichen 87, Germany
- B. Rigg, University of Bradford, School of Colour Chemistry, Bradford 7, England.
- E. Rohner, Bekleidungsphys. Institut/Technische Akademie in Hohenstein, D-1721 Hohenstein ü/Besigheim (Württ.) Germany.

- J. A. J. Rouffs, Institute for Perception Research, Insulindelaan 2, Eindhoven, The Netherlands.
- M. Saltzman, Specialty Chemical Division, Allied Chemical Corporation, Morris Township Center, P. O. Box 70, Morristown, New Jersey 07960, U.S.A.
- J. Schanda, Res. Inst. f. Techn. Phys., Budapest-Ujpest 1. Pf. 76, Budapest, Hungary.
- W. Schramm, CIBA-GEIGY AG, Basel (Rosental), Switzerland.
- W. Schultze, Hauptlaboratorium Badische Anilin- & Soda-Fabrik AG, 67 Ludwigshafen am Rhein, Leuschnerstrasse 42, Germany.
- F. T. Simon, Dept. of Textiles, Clemson University, Clemson, S. C. 29631, U.S.A.
- W. S. Stiles, 89 Richmond Hill Court, Richmond, Surrey, England.
- D. Strocka, Ing. Abt. 18 Bayer, 509 Leverkusen, W. Germany.
- H. Terstiege, Bundesanstalt für Materialprüfung, Fachgruppe Farbmetrik, 1 Berlin 45, Unter den Eichen 87, W-Germany.
- G. Tonnquist, Res. Inst. of National Defense, SS-10450, Stockholm 80, Sweden.
- P. W. Trezona, National Physical Laboratory, Department of Trade and Industry, Teddington, Middlesex, England.
- A. Valberg, Physikalisches Institut, Klingelbergstrasse 82, CH-4000, Basel, Switzerland.
- J. Verschuere-van Helden, c/o Levis N.V., Leuvensesteenweg 167-199, B-1800 Vilvoorde, Belgium.
- J. J. Vos, Institute for Perception TNO, Kampweg 5, Soesterberg, The Netherlands.
- H. G. Wagner, 61 Darmstadt, Ulvenbergstrasse 4, Germany.
- P. L. Walraven, Institute for Perception TNO, Kampweg 5, Soesterberg, The Netherlands.
- Ch. M. M. de Weert, University of Nijmegen, the Netherlands.
- W. D. Wright, Imperial College of Science and Technology, Applied Optics Section, Department of Physics, Prince Consort Road, London, SW7, England.
- G. Wyszecki, National Research Council, Ottawa 8, Ontario, Canada.

CONTENTS

1. W. S. Stiles, Line element in colour theory: A historical review.	1
2. R. M. Boynton and H. G. Wagner, Color differences assessed by the minimally-distinct border method.	26
3. P. W. Trezona, A new method of large field colour matching leading to a more additive metric.	36
4. N. Kambe, Wavelength discrimination measured by chromatic flicker stimuli.	50
5. A. Valberg and T. Holtsmark, Similarity between JND-curves for complementary optimal colours.	58
6. J. J. Vos and P. L. Walraven, A zone-fluctuation line element describing colour discrimination.	69
7. S. L. Guth, A new color model.	82
8. B. H. Crawford, Brightness units and threshold units.	99
9. H. Terstiege, The influence of the fundamental primaries on chromatic adaptation and colour-difference evaluation under different illuminants.	103
10. General discussion I	114
11. T. Indow and K. Ohsumi, Multidimensional mapping of sixty Munsell colors by nonmetric procedure.	124
12. K. Richter, Description of colour attributes and colour differences.	134
13. D. B. Judd, Perceptually uniform spacing of equiluminous colors and the loci of constant hue.	147
14. D. L. MacAdam, Role of luminance increments in small color differences	160
15. G. Wyszecki and G. H. Fielder, New color-matching ellipses.	
16. F. Parra, Continuation of the study of colour thresholds.	188
17. I. Nimeroff, Does the 1964 CIE $U^*V^*W^*$ have a spectrum locus?	193
18. P. Kowaliski, Equivalent luminances and the reproduction of colors.	200
19. F. W. Billmeyer, Jr., E. D. Campbell, G. L. Kandel and J. Mac-Millan, Small and moderate color differences. I. Visual evaluation of FMC-1 and FMC-2 metrics.	211
20. I. G. H. Ishak and S. Roylance, Colour tolerances in the paint industry.	226
21. F. Malkin and A. Dinsdale, Colour discrimination studies in ceramic wall-tiles.	238
22. W. Schultze, The usefulness of colour-difference formulae for fixing	254

colour tolerances.

23. S. M. Jaeckel, Colour-difference formulae for match-acceptability.	266
24. E. Coates, S. Day, J. R. Provost and B. Rigg, Colour-difference equations for setting industrial colour tolerances.	286
25. K. McLaren, Multiple linear regression: A new technique for improving colour difference formulae.	296
26. F. T. Simon, Industrial color tolerances by XI-ETA formulae.	308
27. L. F. C. Friele, FMC-metrics: What next?	315
28. General discussion II	326
29. General discussion III	330
30. G. Wyszecki, Recent developments on color-difference evaluations.	339
31. L. F. C. Friele, A survey of some current color difference formulae.	380
Author index	387

1 THE LINE ELEMENT IN COLOUR THEORY: A HISTORICAL REVIEW

W. S. STILES

*Formerly of the National Physical Laboratory
Teddington, England*

It is very fitting that this Symposium should be associated with the name of Helmholtz. In the second edition of his great work on physiological optics (Helmholtz, 1896) he devoted a considerable section to the notion of a line element in colour space. The original idea he had developed a few years earlier in papers to the *Journal for the Physiology and Psychology of the Sense Organs* (Helmholtz, 1891, 1892). The relevant volume—Volume 2—of the third edition of the treatise did not appear until 1911, seventeen years after Helmholtz's death, and the distinguished editors saw fit to omit the passages on the line-element. This no doubt was a pity for the development of ideas on line-elements, and perhaps also serves as a reminder to all editors of the hazards of pruning great works.

Helmholtz wished to extend the concept of brightness sensitivity contained in the Weber-Fechner psychophysical law, to the more complicated case of colour sensitivity. In his discussion of the discrimination of brightness differences (ΔB) he dealt at length with the deviations from the simple Weber principle:

$$\text{Weber fraction} = \frac{\Delta B}{B} = \text{constant},$$

that occur at low and very high brightnesses B . At low levels he accepted that the increasing values of the fraction arose from the existence in the eye of so-called self-light (*Eigenlicht*) unrelated to the external stimulus and associated in his view with the luminous appearances seen by an observer kept in the dark. It is interesting that Helmholtz was intrigued by the fluctuating, patchy and granular appearance of the self-light—a premonition, one might almost say, of current concern with noise and quantum fluctuations in the visual system. The effect on the Weber fraction of an average level B_0 of self-light was simply allowed for to give:

$$\frac{\Delta B}{(B+B_0)} = \text{constant.}$$

Helmholtz's calculations on the unevenness of the self-light led him to refine this expression by the addition of a further term to obtain:

$$\frac{\Delta B}{B+B_1} \left\{ 1 + \frac{A}{(B+B_1)(B+B_2)} \right\} = \text{constant}$$

where the three constants A , B_1 , B_2 replacing the single constant B_0 depend on some simple assumptions about the spatial and intensity variations of the self-light as well as its average value. Tests of this refinement using König and Brodhun's (1889) classical data on brightness discrimination for the case of monochromatic red light were not very satisfactory and Helmholtz reverted to the simpler approximation in his later treatment of the line-element.

At the other extreme of very high brightness, Helmholtz corrected the psychophysical law for an increase in the Weber fraction produced by what he described as the glare of the field. Basing himself again on the discrimination data of König and Brodhun—which certainly show for all colours an increase in the Weber fraction at brightnesses above some 50,000 trolands—he introduced into the law an additional factor

$$\left(\frac{1}{1+gB} \right)$$

where g is a small constant, the same for all colours. This gave:

$$\frac{\Delta B}{B+B_0} \cdot \frac{1}{1+gB} = \text{constant.}$$

Much doubt has been thrown by later investigators on the upturn in the Weber fraction at high intensities. Although we now know that the Weber fraction of the scotopic or rod mechanism begins to increase at about 100 scotopic trolands and ascends to very high, so-called saturation values in the region of 1000-3000 scotopic trolands, and it has also been shown, quite recently, by Alpern, Rushton and Torii (1970) that the cone mechanisms exhibit at higher levels a similar approach to saturation, the phenomenon for cones is not observable when the eye is fully adapted to a steady field. In that case the Weber fraction of cone vision remains constant for indefinitely high brightnesses. In the late twenties, John Guild, a co-founder of the CIE system, liked to show an experiment in which a tungsten filament was focussed on the observer's eye and by suitable transparencies weak contrast patterns were introduced into the very bright Maxwellian field so obtained. Initially the observer saw merely a dazzling bright area but after a few minutes steady gazing he saw the patterns as well as at much lower brightness levels. It may fairly be said that, provided

the observer is fully adapted to a steady uniform field of adequate size, the high intensity upturn in the Weber fraction does not occur, at least until brightnesses much higher than 50,000 trolands and approaching pathological intensity, are reached. To get back to Helmholtz, he formally included the glare correction in his line-element, but in fact it played no significant part in the applications he studied.

To switch from brightness to colour discrimination Helmholtz based himself on Young's trichromatic theory in which any colour can be thought of as containing definite quantities of three uniquely specified fundamental colours. Helmholtz's argument—put forward just as probable hypothesis to be justified by its results—ran briefly as follows. The discrimination of any two colours must depend on the possibility of discriminating the differences in the quantities of the respective fundamentals they contain. The magnitude or conspicuousness—*Deutlichkeit*—of the sensation difference ds between two colours which are only slightly different is some function of the magnitudes of the sensation differences ds_1 , ds_2 , ds_3 of the respective fundamentals. ds can be zero only if ds_1 , ds_2 , ds_3 are all zero because a sensation difference in one fundamental cannot cancel out a sensation difference in another. If higher powers of the small quantities ds_1 , ds_2 , ds_3 can be ignored, the simplest function with the property just mentioned is a positive definite quadratic form in ds_1 , ds_2 , ds_3 . But, assuming physiological independence of the three fundamental processes involved, it is reasonable to expect that the resultant conspicuousness ds will be the same if one of the original colours contains more of all three fundamentals than the other, or if, alternatively, one of the original colours contains more of one fundamental and less of the other two. Thus the signs of the quantities ds_1 , ds_2 , ds_3 cannot matter, and the cross-product terms in the quadratic form for ds^2 must all have zero coefficients; as ds^2 must also reduce to ds_1^2 if $ds_2 = ds_3 = 0$, and so on, the expression for ds^2 takes the simple form:

$$ds^2 = ds_1^2 + ds_2^2 + ds_3^2$$

For each of the sensation differences of the fundamentals, ds_1 , ds_2 , ds_3 , Helmholtz assumed the expression given by the Weber-Fechner Law for brightness discrimination with the quantity of the fundamental replacing brightness. If we drop the glare factor and the term for unevenness of the self-light—Helmholtz very soon drops them in his own analysis—we obtain the famous Helmholtz line-element:

$$ds^2 = \frac{dR^2}{(R+a)^2} + \frac{dG^2}{(G+b)^2} + \frac{dB^2}{(B+c)^2}$$

The quantities of the fundamentals— R , G , B —must of course be evaluated using a unique set of fundamental spectral sensitivities which, initially, Helmholtz left undetermined. The three terms differ otherwise only in the magnitudes

a, b, c of the self-light in the respective fundamentals, a difference that becomes increasingly unimportant as the level is raised. At such high levels, the element predicts that for pure brightness discrimination, that is, when

$$\frac{\Delta R}{R} = \frac{\Delta G}{G} = \frac{\Delta B}{B},$$

the Weber fraction will have the same value for all colours, in accord with König and Brodhun's experimental findings.

The first—and a highly significant—application of the line-element made by Helmholtz was to the determination of pairs of neighbouring colours of "greatest similarity". He envisaged a comparison of one colour having fixed chromaticity and fixed intensity with another of slightly different chromaticity and of variable intensity. The line-element then enables the intensity of the second colour to be calculated at which the quantity ds^2 is minimal, the two colours then being said to have greatest similarity. This procedure corresponds closely with what is done experimentally in determining a relative luminous efficiency curve by the step-by-step method, and neighbouring colours of greatest similarity could also be described as colours of the same brightness or, more precisely, of the same "small-step" brightness. The condition for equal small-step brightness of neighbouring colours and the value of ds^2 when this condition is satisfied, may be written in the case of the Helmholtz element:

$$0 = \frac{dR}{R} + \frac{dG}{G} + \frac{dB}{B}$$

and

$$ds^2 = \frac{k^2}{3} \left[\left(\frac{dR}{R} - \frac{dG}{G} \right)^2 + \left(\frac{dG}{G} - \frac{dB}{B} \right)^2 + \left(\frac{dB}{B} - \frac{dR}{R} \right)^2 \right]$$

where an intensity level high enough for the self-light in all fundamentals to be negligible, is assumed.

These results were needed by Helmholtz in his explanation of König and Dieterici's (1884) measurements of the smallest perceptible difference of wavelength through the spectrum when the intensities of the monochromatic stimuli being compared were adjusted to give greatest similarity, i.e. equal small-step brightness. The observed difference of wavelength $\Delta\lambda$ should be given by the above expression for ds^2 on inserting for R, G, B and their variations values appropriate to the spectral colours, and assuming for ds a small fixed value.

$$\frac{ds^2}{k^2} = \frac{(\Delta\lambda)^2}{3} \left[\left(\frac{1}{R} \frac{dR}{d\lambda} - \frac{1}{G} \frac{dG}{d\lambda} \right)^2 + \text{similar terms in } (G, B) \text{ and } (B, R) \right]$$

Helmholtz soon found that his theory broke down if he used as fundamentals the so-called "Grundempfindungen" curves selected by König largely from a

consideration of the colour-matching properties of dichromats. But by transforming linearly to a new set of fundamentals—his so-called “Elementarfarben” or “Urfarben”—he obtained fair agreement with König and Dieterici’s data. The Elementarfarben, each with a broad spectral sensitivity curve exhibiting two well-marked maxima, were very unlike the König Grundempfindungen. The position of the spectrum locus in a colour triangle with the Elementarfarben at the corners, shown in Fig. 1 due to Helmholtz, emphasises

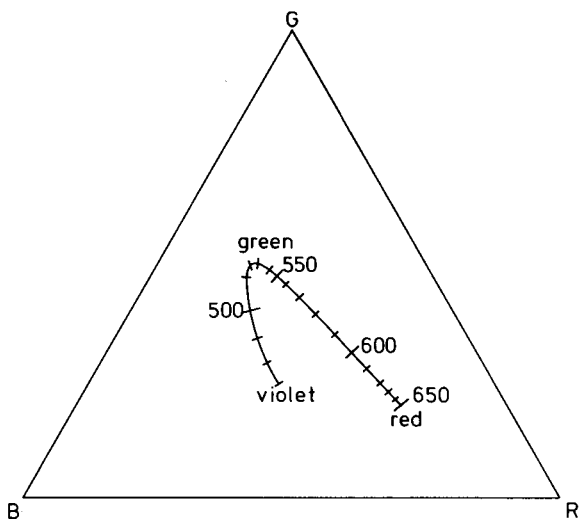


Fig. 1 The position of the spectral locus in Helmholtz' colour triangle. In the corners Helmholtz' "Elementarfarben" (Helmholtz, 1892)

the big difference between his fundamentals and sets that would be acceptable to-day. If true, the Elementarfarben would entail the abandonment of the idea that dichromatic vision arose simply by the absence of one fundamental. Helmholtz seems to have rated his Elementarfarben highly and contemplates explanations of dichromasy on the lines of Fick's principle of fused fundamentals, which would entirely avoid the mere assumption of a lacking fundamental.

Another concept in colour-metrics introduced by Helmholtz was that of the "shortest colour line". If any line joining two points in three-dimensional colour space is marked out in steps of just perceptible difference as defined by the line-element with ds put equal to a small constant, the number of steps depends on the path of the line, and is minimal for a particular path—the shortest colour line or geodesic. For Helmholtz's element the geodesics are readily obtained; in a transformed space in which the coordinates are the logarithms of the tristimulus values plus in each case the appropriate self-light constant— $\log(R+a)$, $\log(G+b)$, $\log(B+c)$ —the geodesic between two points is simply the straight line joining them. One conclusion was that for just one

colour quality—the one in which the tristimulus values are in the same proportion as the corresponding self-lights—change of intensity could produce no change in apparent hue. Helmholtz recognised yellowish-white as having experimentally this property, all other lights tending to this colour as their intensities are made very high. In his discussion of colour appearances at lower levels, in particular, in considering the Bezold-Brücke effect the self-light constants again play an important part.

Thus over eighty years ago we had the concepts of a line-element, of maximal similarity or small-step brightness-matching of neighbouring colours, and of geodesics in colour space, all due to Helmholtz. But it was not until some thirty years later that these concepts were critically re-examined, and it was to another great mathematical physicist Erwin Schrödinger that we owed the next advances. In three masterly papers under the title “A Theory of the Colour Metric in Photopic Vision” Schrödinger (1920) deployed in concise mathematical form the trichromatic system as then current. The last of these papers devoted to higher colour metrics contains his work on line-elements.

Like Helmholtz, Schrödinger started with a general quadratic form for the line-element:

$$ds^2 = \sum_{i,j=1}^3 g_{ij} dX_i dX_j$$

where the tristimulus values X_1, X_2, X_3 are not necessarily based on fundamentals. But unlike Helmholtz, he did not commit himself to the view that the resultant discrimination depends on discrimination by three independent processes each involving just one fundamental. The Helmholtz condition for small-step brightness match between neighbouring colours becomes for the unrestricted element

$$0 = \sum_{i,j=1}^3 g_{ij} X_i dX_j$$

and Schrödinger points out that such a condition will not generate consistent surfaces of colours all in small-step brightness with their neighbours unless the coefficients g_{ij} satisfy a certain mathematical condition which need not be quoted. The Helmholtz element certainly satisfies this condition but for not too low brightness levels it leads to equibright surfaces of the form:

$$R.G.B. = \text{constant}.$$

It also implies for the small-step relative luminous efficiency a curve which, with any acceptable choice of fundamentals, is completely different from the experimental V_λ curve. With the Grundempfindungen or Elementarfarben as fundamentals the computed V_λ has two maxima, with other currently acceptable sets it certainly has just one maximum but is impossibly broad, and

only for the Hecht (1932) fundamentals each of which approximates to the observed V_λ curve would the Helmholtz derivation give approximate agreement. Obviously, according to the Helmholtz element, small-step brightness is not an additive property.

To meet all his criticisms of the Helmholtz element, Schrödinger proposed a new form:

$$ds^2 = \frac{1}{l_R R + l_G G + l_B B} \left\{ \frac{l_R dR^2}{R} + \frac{l_G dG^2}{G} + \frac{l_B dB^2}{B} \right\}$$

where the tristimulus values R, G, B are to be evaluated on the basis of the König Grundempfindungen, which are not radically different from modern sets of acceptable fundamentals. It will be noted that there are no cross-product terms— $dR.dG$, etc.—in the expression, so that this element again implies three mechanisms that contribute independently to the resultant ds . But now the magnitude of each mechanism's contribution depends on all three tristimulus values. It is readily shown that the Schrödinger element leads to surfaces of equal small-step brightness that are the planes— $l_R R + l_G G + l_B B = \text{const.}$; it makes brightness discrimination the same for all colours, and it also makes small-step brightness an additive property. It predicts tolerably well—although Schrödinger expressed disappointment at not getting better agreement—experimental data on relative luminous efficiency and on wavelength discrimination.

Schrödinger elaborated to some extent the idea of geodesics in colour space. He took the value of the integral of ds along the geodesic between two colours as representing their "dissimilarity", and defined heterochromatic matching between widely different colours as the minimisation of their dissimilarity when the intensity of one colour is kept fixed while that of the other is varied. For his element, as for that of Helmholtz, this definition leads to the same equibrightness surfaces as small-step brightness-matching. Starting in various directions from any colour of fixed chromaticity and brightness, a network of geodesic lines can be drawn which map out the whole of colour space. In particular, starting from an achromatic white, geodesic lines terminating on the spectral colours and purples can be drawn along which brightness is constant. Schrödinger made the plausible assumption that the apparent hue on these lines would also be invariable and obtained a theoretical picture of the change of hue with saturation already measured directly by Abney (1913) and evidenced of course in the Munsell colour system.

Schrödinger's element was not designed to apply to very low brightness levels and it contains no constants corresponding to self-light. He noted that when these were included—which is easily done—equibrightness surfaces cease to be planes and brightness is no longer additive, although these effects are only significant at low levels.

In the years following Schrödinger's considerable contribution to line-element theory, some not very successful attempts were made to validate his element experimentally. I think the next significant step was the adoption by the *CIE* (1931) of an agreed colorimetric system for the precise specification of the tristimulus values of any stimulus. Practical applications of the system made it increasingly desirable to know for pairs of stimuli located in different regions of colour space, the differences in their tristimulus values corresponding to a just perceptible difference, and particularly to know this for pairs of colours of the same luminance. Outstanding was the paper by Judd (1935) in which he collected all the data he could find on just discriminable differences of hue and saturation at constant brightness, and looked for a linear transformation of the *CIE* tristimulus values to another equivalent system which for a plane of constant luminance would make just perceptible differences correspond to equal displacements of the representative colour points. For the unmodified *CIE* system such a correspondence is not even roughly true. Neither the Helmholtz nor the Schrödinger line-elements would in fact predict that Judd's objective could be realised. However, Judd showed that a transformation was possible which, while falling a good deal short of perfection, gave an approximation of great practical interest. Judd's work was soon followed by further suggestions for uniform chromaticity scales, all obtained by transformations of the *CIE* system and usually defined for uniform steps in the x-y chromaticity diagram rather than in the constant luminance colour triangle of Judd's original system. From the standpoint of the line-element, an important development was the spur this work on uniform chromaticity scales gave to the procurement of more satisfactory and more comprehensive data on colour discrimination, and two major investigations with this aim appeared in 1941 and 1942 respectively. Wright (1941) determined, by direct observations, just noticeable differences along straight lines forming a network in a plane of constant brightness, the observer himself equalising the brightness and adjusting the colour difference of the two test patches. MacAdam (1942), on the other hand, working with colours in a plane of constant *CIE* luminance ($Y = \text{constant}$) determined the spread of colour matches in the chromaticity diagram for a set of 25 test colours, and derived the famous MacAdam ellipses representing this spread. We must also note in about the same period the systematic determination of the *CIE* tristimulus values for the samples in the Munsell colour system (Nickerson and Newhall, 1943) which embodies much psychophysical material on loci of constant hue and saturation and on uniform scaling in the colour solid. But MacAdam's pioneer research, followed by others by Brown and MacAdam (1949) and by Brown (1952), extending the measurements to three-dimensional matches and leading to discrimination ellipsoids, is the most directly relevant to the line-element.

MacAdam adopted what I would call the strictly empirical approach. He

and Brown were aiming at the facts of colour discrimination unalloyed by theoretical pre-occupations and interpretations in terms of visual colour mechanisms. In connection with this chain of researches a number of points, partly methodological, may be mentioned. In the matches whose precision was being studied, it was important that the two stimuli should not only approach identity of colour but also identity of spectral composition, and the experimental set up was always arranged to achieve this end. For comparison with other results some knowledge of the relation of matching errors to least perceptible differences, was required. The precise relation is difficult to define in a theoretical manner and MacAdam's experimental determination of the approximate factor for converting standard deviations of matches into least perceptible differences, provided the necessary link. The mathematician Silberstein (1945), and MacAdam probed the question whether repeated colour matches corresponded to a multivariate normal distribution which would then justify the use of ellipses and ellipsoids to represent the matching spread. Silberstein (1943 and earlier) was much interested in the intrinsic properties of the colour domain, and his examination of the shape and Gaussian curvature of a surface, embedded in three-dimensional space, on which the suitably transformed MacAdam ellipses would plot as circles of equal size, led MacAdam (1944) to make a model of this ideal, if rather contorted, uniform chromaticity surface. His striking model is a kind of landmark in the subject, and I hope that somewhere it is preserved.*

When looking back on much of the elegant mathematical theory expended on colour metrics, I can't help recalling a cutting remark of the late Dr. P. J. Bouma in his fascinating book on colour published in 1947 just before he died. Commenting on some of Silberstein's contributions, he remarks "In his speculations, the fundamental assumptions are so concealed under the mathematical brushwood that it is difficult to make out whether the considerations have any theoretical value. They certainly have no practical significance" Bouma goes on to give references to others who have contributed to the mathematical undergrowth—and I note that my own name occurs in the list.

To get back to the MacAdam-Brown data—the value of which was of course not in question—we find that applications to practical colour tolerance problems entailed a by no means straightforward interpolation and extrapolation from data obtained at a necessarily limited number of colour points. This was undertaken by MacAdam and by others to provide adequate working systems and, in effect, to define throughout colour-space purely empirical line-elements for the two and three dimensional cases:

* Interruption MacAdam: yes, it is.

In tristimulus space:

$$ds^2 = \sum_{i,j=1}^3 k_{ij} X_i X_j$$

$$(X_1 = X_{CIE}, X_2 = Y_{CIE}, X_3 = Z_{CIE})$$

In equiluminance plane:

$$ds^2 = g_{11} dx^2 + 2 g_{12} dx dy + g_{22} dy^2$$

$$(x, y = CIE \text{ chromaticities})$$

In chromaticity/log luminance space:

$$ds^2 = \sum_{i,j=1}^3 p_{ij} dw_i dw_j$$

$$(w_1 = x, w_2 = y, w_3 = \frac{1}{3} \log_{10} Y)$$

The last form is convenient in displaying the effects of luminance level; its coefficients p_{ij} would be independent of Y or w_3 if the luminance contrast sensitivity and the equiluminance ellipses did not vary with luminance. In general the coefficients— k_{ij} , g_{ij} , p_{ij} —are empirical functions of the complete sets of variables concerned.

It is appropriate to mention here a succession of colour-difference formulae combining both luminance or lightness, and colour differences, such as the Adams chromatic valence formula, and the expressions put forward by Saunderson and Milner, Scofield, Hunter, Reilly and colleagues, and others (as referred by Wyszecki and Stiles, 1967), resting in most cases on the Munsell system. Generally, total colour difference is represented as the square root of the sum of three squared difference terms, and in each case what is provided is essentially a form of line-element. Much thought on colour theory went into many of these elements, and in mentioning them in the context of the empirical elements I have regard to their main preoccupation with the practical problems of measuring colour differences.

A rather different approach to colour metrics is linked with the idea that a line-element should in some sense be derived from the knowledge we have of the visual processes that make colour discrimination possible. It should draw on concepts and experiments that go beyond the direct measurements of colour discrimination and should provide a model that would not only reproduce tolerably well existing blocks of colour discrimination data but by its structure and by the identification of its parameters, should display the connection of discrimination with basic attributes of visual mechanisms. Ideally it should provide a framework capable of modifications and extensions which may help in understanding visual behaviour in more diverse observational situations. I have used the term "inductive" to describe line-elements where these considerations have operated. Naturally there is no sharp distinction

between the empirical and inductive approaches. As we learn more of visual processes we may expect this to be reflected in more sophisticated inductive line-elements. Such development has been evident I think in the history of the line-element in the last decade or so, but a simpler and earlier example is provided by the modification of the Helmholtz element proposed by Stiles (1946). It arose from investigations which in fact had nothing to do with colour-matching and were not specifically concerned with colour discrimination. The experiments consisted in determining the threshold for detection of a small test stimulus of one colour applied as an additional stimulus to a steady and uniform adapting field usually of a different colour (two-colour threshold technique). From the variations of the increment threshold, so measured, with the three variables: adapting field intensity ranging from zero to very high levels, and the colours, usually monochromatic specified by wavelength, of test and field, it was concluded that the observed increment could be regarded as the resultant—roughly equal to the lowest—of component increment thresholds each associated with a different response mechanism. The results yielded for each mechanism its characteristic relative spectral sensitivity curve and, for foveal rod-free vision of the test stimulus, three mechanisms were distinguished; blue-sensitive, green-sensitive and red-sensitive. Although not dependent on any colour-matching data and not of the same precision as such measurements, the sensitivity curves of the hypothetical mechanisms could be represented approximately by specific linear combinations of the *CIE* colour-matching functions, and were generally similar to König's Grundempfindungen. Most striking however were the values obtained for the limiting Weber-Fechner fractions of the component mechanisms—that is to say, the ratio of the mechanism's increment threshold to the field intensity when test and field had the same colour and when the field intensity was high enough for the region of constant Weber-Fechner fraction to be reached. The blue-sensitive mechanism was found to have a limiting Weber-Fechner fraction about 4.5 times that of the green-sensitive, the value for the red-sensitive mechanism being slightly lower, by a factor of about 0.8, than for the green-sensitive.

The injection of these findings into a line-element of the Helmholtz type led to a proposed modified form:

$$ds^2 = \left[\frac{dR'}{\rho} \zeta(R') \right]^2 + \left[\frac{dG'}{\gamma} \zeta(G') \right]^2 + \left[\frac{dB'}{\beta} \zeta(B') \right]^2$$

$$\rho : \gamma : \beta = 0.8 : 1 : 4.5$$

The empirical function $\zeta(\dots)$ defines the shape of the curve showing the increase of threshold with field intensity—which was found to be practically the same for all mechanisms. It didn't differ much from the shape corresponding to the Weber-Fechner psychophysical law including self-light, and if the difference is ignored the modified element becomes:

$$ds^2 = \left[\frac{1}{\rho} \cdot \frac{dR}{R + R_0} \right]^2 + \left[\frac{1}{\gamma} \cdot \frac{dG}{G + G_0} \right]^2 + \left[\frac{1}{\beta} \cdot \frac{dB}{B + B_0} \right]^2$$

where R, G, B are the tristimulus values based on the relative spectral sensitivity curves of the mechanisms (using best CIE combinations), ρ, γ, β are proportional to their limiting Weber-Fechner fractions, and R_0, G_0, B_0 are the respective self-light constants, all these quantities being provided by the increment threshold measurements.

The small-step brightness for the modified element, derived by the Helmholtz method, is still non-additive, although less grossly so than before. On the other hand, the derived relative luminous efficiency curve agrees well with experimental small-step V_λ curves such as those measured by Gibson and Tyndall (1923). As regards the additivity failure, there is in fact hardly any experimental evidence on this aspect of small-step brightness. Flicker brightness-matching is certainly very nearly additive but the flicker measurement is of a quite special kind to which the line-element cannot be applied without auxiliary assumptions. Direct comparison heterochromatic brightness-matching shows serious breakdowns in additivity when colours of widely different chromaticity are involved. However, at least some of these discrepancies are in the opposite sense to what would be predicted by the modified line-element using Schrödinger's geodesic definition of brightness-matching. The criterion that for all colours the Weber-Fechner fraction should be the same, except at very low levels, is certainly satisfied. Also, for wavelength discrimination by normal trichromats there is fair agreement between predictions and experimental data. Quite recently Trabka (1968) has shown that a similar agreement is obtained if the element is applied to protanopic and tritanopic vision, by omitting one mechanism, and to deuteranopic vision, by allowing an effective fusion of two mechanisms. Calculations of the 25 MacAdam ellipses however, while producing a generally similar set with fair correlation of computed and observed ellipse orientation and area, showed no correlation for ellipse eccentricity. That a deep-rooted discrepancy is involved is further suggested by the fact that the calculated analogue of MacAdam's famous uniform chromaticity surface has quite different properties of Gaussian curvature.

The discrepancy is associated, in part at least, with the absence from the line-element of cross-product terms— $dR.dG$., etc. For any element without such terms i.e. of the form:

$$ds^2 = [f_R(R, G, B).dR]^2 + [f_G(R, G, B).dG]^2 + [f_B(R, G, B).dB]^2$$

— which includes Schrödinger's element—a certain type of independence of the component mechanisms is indicated. Some ten years ago, I tested this independence, and the validity of a line-element of the form just given, using an increment threshold method, in work later continued by Boynton, Ikeda and

Stiles (1964) (See also Stiles, 1967). With the eye adapted to a uniform field of fixed intensity and colour, the increment threshold was measured for two monochromatic test stimuli, first separately and then for a whole range of mixtures of the two wavelengths in various fixed proportions. If the two wavelengths are sufficiently close together it is clear that the threshold for any mixture must correspond very nearly to linear summation, that is to say, if the quantity of each wavelength in a mixture is expressed in units of the threshold for that wavelength acting alone, then at the mixture threshold these two quantities must add up to unity. On the other hand if the wavelengths are widely separated, so that, for example, the threshold for one wavelength is determined predominantly by the dR^2 term in the element, while that for the other wavelength is determined mainly by one of the other terms, then the sum of squares form of the line-element shows that linear summation will not hold, and, in fact, a more intense mixture will be needed to reach mixture threshold. The two wavelengths then sum their effects less completely. But the form of the element also shows that this reduced summation is of strictly limited extent. A summation index can be derived from the experiments with any two wavelengths which has the value $\log 2 = 0.30$ for linear summation and which for an element with no cross-product terms can never drop below $\log \sqrt{2} = 0.15$. As the example in Fig. 2 shows this is not borne out by experiment, which in

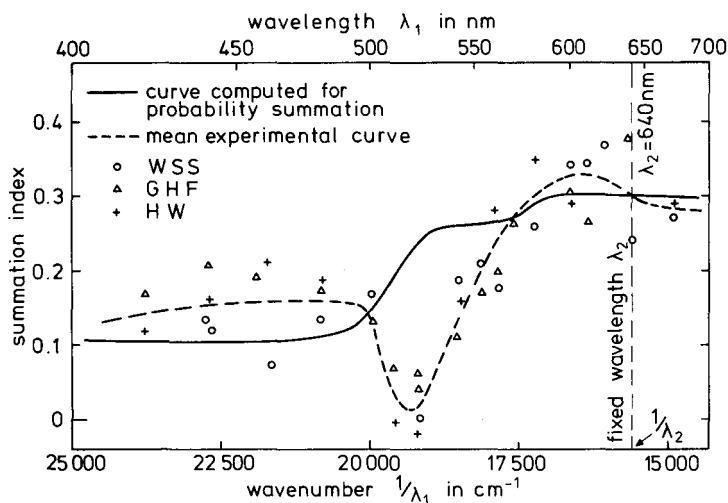


Fig. 2 Summation index as a function of wavenumber for three observers. λ_1 is varied, λ_2 is kept constant at 640 nm (Stiles, 1967).

this case refers to a high intensity green adapting field. For one wavelength fixed in the red (640 nm), and the other variable through the spectrum, the summation index conforms tolerably well with expectation when the second

wavelength is in the red or in the blue, but around 530 nm it drops anomalously to zero. This and other related anomalies under a variety of conditions, including cases where a positive increment of one wavelength is mixed with a negative increment of the other, provided strong evidence that the line-element must include cross-product terms in the differentials dR , dG , dB . The mixture experiments represent a fairly searching test, and the summation anomalies are sufficiently small to have escaped detection in the earlier increment threshold studies. It seems certain however that the modified line-element must be regarded as only a first approximation. The results are complicated and, as yet, no satisfactory generalisation of the modified element including cross-product terms has been reached along these lines.

Mention should be made of very similar mixture experiments by Guth *et al* (1968) for the special case of zero adapting field. These were directed, at least initially, to the more primitive question whether, at absolute threshold, linear summation, or Abney's additivity law for brightness to use Guth's phraseology, would hold for any two wavelengths. This would not be expected for any hitherto suggested line-element, and Guth found that it did not in fact occur. He has carried his analysis of the zero field results a good deal further and has made interesting comparisons with non-additivity in direct comparison heterochromatic brightness-matching (Guth *et al.*, 1969).

A deeper interpretation of the empirical function $\zeta(\dots)$ relating the increment threshold of a mechanism to the field intensity was made by Trabka (1968). He thought of each mechanism as a set of similar radiation detectors with discriminating power defined by the signal to noise ratio at output. If the average rate of absorption of photons at input is m , and the stimuli to be discriminated correspond to a difference dm , then following the earlier theories of De Vries (1943) and of Rose (1948) he equated the noise at input to the statistical fluctuation appropriate to a Poisson distribution — $\sqrt{m}$ — and obtained for ideal detectors:

$$\left(\frac{\text{SIGNAL}}{\text{NOISE}}\right)_{\text{INPUT}}^2 = \frac{dm^2}{m} = \left(\frac{\text{SIGNAL}}{\text{NOISE}}\right)_{\text{OUTPUT}}^2$$

However this leads to a law of increase of threshold with field that, taken over the whole range from zero to high levels, is seriously at variance with experiment. He derived a modified expression for the signal to noise ratio by replacing the ideal detectors by imperfect ones which add to m an average (fluctuating) dark current m_0 , and for which there is a dead-time τ after the absorption of each photon producing output, during which no further absorption will contribute to the output. These assumptions give:

$$\left(\frac{S}{N}\right)_{\text{OUTPUT}}^2 = \frac{\tau^{-1} dm^2}{m^2 + 2bm + c}$$

with

$$b = (1 + 2m_0\tau)/2\tau, \quad c = m_0(1 + m_0\tau)$$

and

$$\zeta(R) = \frac{1}{(p_2 R^2 + p_1 R + p_0)^{\frac{1}{2}}}$$

(p_2, p_1, p_0 constants dependent on m_0, τ)

The dark-current is of course a near cousin of Helmholtz's self-light, and the concept of dead-time has been used by several investigators in this connection, notably by Van der Velden (1949). With suitably chosen values for the constants in the above expression, Trabka found he could fit rather well the empirical threshold versus field function $\zeta(\dots)$ used in the modified line-element. In re-developing the latter in a more up-to-date mathematical form Trabka (1969) indicated a possible statistical refinement that might bring in cross-product terms, but this has not I think been followed up.

Before looking briefly at more recent line-elements I have a couple of remarks to make on general points. Firstly, although the number of component mechanisms concerned in foveal rod-free discrimination is usually taken as three, leading to three terms in the element, this is not strictly necessary. It is conceivable that certain neural systems in retina and brain are specialised for rather limited perceptual tasks, associated perhaps with a particular colour or colour contrast, or, again, at different levels of stimulation, apparently similar discriminations may be handled by different mechanisms. There is already some evidence that there may be more than one blue-sensitive component mechanism, and that the detection of a short wavelength incremental test stimulus on a long wavelength field may be effected through one or other mechanism depending on the intensity level. The inclusion of more than three terms in a foveal line-element need not in any way violate tristimulus specification nor the broad organisation of our colour sensations under the concepts of hue, saturation and brightness. Zonal theories of colour vision also contemplate more than three mechanisms but in a rather different sense.

My second point concerns the way in which a particular mechanism contributes to the line-element. In many, although perhaps not in all cases, the contribution may be thought of as determined by the ratio ξ of the difference of the mechanism's response F to the neighbouring stimuli divided by an effective threshold difference T appropriate to all the prevailing visual conditions. In particular interpretations, this ratio can be identified as a signal to noise ratio. The response and threshold functions will depend in general on the fundamental tristimulus values R, G, B in the discrimination field, and also on the surround stimuli, adaptation etc. Excluding the latter complications the contribution of the i^{th} mechanism is represented by

$$\xi_i = \frac{\Delta F_i}{T_i} = \frac{\frac{\partial F_i}{\partial R} \cdot \Delta R + \frac{\partial F_i}{\partial G} \cdot \Delta G + \frac{\partial F_i}{\partial B} \cdot \Delta B}{T_i(R, G, B)}$$

How are the contributions of different mechanisms to be combined? From the standpoint of actual threshold measurements the quantity ξ_i determines the chance of detecting the difference of neighbouring stimuli by the action of the i^{th} mechanism acting alone, and T_i in that case represents the value of ΔF_i for some agreed threshold chance, such as 50 per cent. At threshold ξ_i is then equal to unity. The resultant threshold when several mechanisms contribute independently to detection, is then derivable from the curves for the several mechanisms that relate the probability of detection with the value of ξ_i . As was emphasised, for example, by LeGrand in his 1948 book on Physiological Optics, this approach does not necessarily lead to a threshold condition in the sum of squares form:

$$1 = \xi_1^2 + \xi_2^2 + \dots$$

The form and steepness of the probability of detection curves must be a factor. However although the sum of squares form may not be quite right, its use is unlikely to obscure other salient features of various line-elements. On the other hand, for a line-element in that form there is an implication that the mechanisms corresponding to the respective terms are contributing independently to discrimination.

Insertion of the expansions for the contributions ξ_i into the sum of squares condition shows at once that for an element containing no cross-product terms in $dR.dG$, etc, the response of every mechanism concerned must depend on one tristimulus value only, although the thresholds may depend on all three. The mechanisms of the earlier elements were of this type. Among mechanisms with response functions depending on more than one tristimulus value, two groups are of special interest in line-element theory: colour-differencing and pure brightness or luminance mechanisms. As regards the former there is some physiological plausibility for the belief that after light has been absorbed in the visual pigment a non-linear step supervenes and it is only after this step that the differencing process occurs. The commonly made assumption of a logarithmic step leads to the following response function and element contribution for a general type of colour-differencing mechanism, based in this case on the red and green tristimulus values

Response: $F_{RG} = \alpha_1 \log_e (R + a_1) - \beta_1 \log_e (G + b_1)$

Threshold: T_{RG}

$$\xi_{RG}^2 = \left(\frac{\Delta F_{RG}}{T_{RG}} \right)^2 = \frac{1}{T_{RG}^2} \left\{ \alpha_1 \frac{\Delta R}{R + a_1} - \beta_1 \frac{\Delta G}{G + b_1} \right\}^2$$

Here $\alpha_1, \beta_1, a_1, b_1$ are constants, the latter pair being usually small self-light constants. The threshold T_{RG} may be a constant or a function of R, G, B ; if it should reduce to a function of the response F_{RG} , this would be an additional way in which the behaviour of the mechanism is self-contained or independent.

With luminance determined by the usual linear form in R, G, B , the response function of a pure brightness or luminance mechanism must be some function of L . Thus for a mechanism of this type the assumption of a non-linear step immediately after light absorption is abandoned. It is however still convenient to use a response function in logarithmic form, to obtain

$$\begin{aligned}\text{Response: } F_L &= \alpha_3 \log_e (l_R R + l_G G + l_B B + l_0) \\ &= \alpha_3 \log_e (L + l_0)\end{aligned}$$

Threshold: T_L

$$\xi_L^2 = \left(\frac{\Delta F_L}{T_L} \right)^2 = \frac{d_3^2}{T_L^2} \left\{ \frac{l_R \Delta R + l_G \Delta G + l_B \Delta B}{l_R R + l_G G + l_B B + l_0} \right\}^2 = \frac{d_3^2}{T_L^2} \left\{ \frac{\Delta L}{L + l_0} \right\}^2$$

It will be observed that the same element contribution can result from different response functions, provided the threshold function is modified in a complementary way. There are of course mathematical conditions controlling this flexibility but they are not needed in the present context.

A complete element made up of two colour-differencing terms—say, the red-green and an analogous green-blue—and a luminance term, of the forms just indicated is obtained by adding the contributions in a sum of squares equation:

$$1 = \xi_{RG}^2 + \xi_{GB}^2 + \xi_L^2$$

Two questions suggest themselves at once. Firstly, at higher levels where the self-light constants have negligible effects, is the ordinary Weber-Fechner fraction for brightness discrimination,

$$H = \frac{\Delta R}{R} = \frac{\Delta G}{G} = \frac{\Delta B}{B},$$

independent of colour? H is given by

$$H^2 = \left[\frac{1}{T_{RG}^2} (\alpha_1 - \beta_1)^2 + \frac{1}{T_{GB}^2} (\alpha_2 - \beta_2)^2 + \frac{\alpha_3^2}{T_L^2} \right]^{-1}$$

and it is clear that if $\alpha_1 = \beta_1$, $\alpha_2 = \beta_2$ and T_L is a constant or depends only on the luminance L , H will be the same for all colours. This will still be true if $\alpha_1 \neq \beta_1$ (or $\alpha_2 \neq \beta_2$) given that T_{RG} (or T_{GB}) depends only on luminance L . The second question is whether the Helmholtz condition for small-step brightness match will generate consistent equibright surfaces. Applying the condition to the present element, again ignoring self-light constants, gives

$$\begin{aligned}0 &= \frac{1}{T_{RG}^2} \left(\alpha_1 \frac{\Delta R}{R} - \beta_1 \frac{\Delta G}{G} \right) (\alpha_1 - \beta_1) + \frac{1}{T_{GB}^2} \left(\alpha_2 \frac{\Delta G}{G} - \beta_2 \frac{\Delta B}{B} \right) (\alpha_2 - \beta_2) + \frac{\alpha_3^2}{T_L^2} \cdot \frac{\Delta L}{L} = \\ &= \frac{1}{T_{RG}^2} \cdot \Delta F_{RG} \cdot (\alpha_1 - \beta_1) + \frac{1}{T_{GB}^2} \cdot \Delta F_{GB} \cdot (\alpha_2 - \beta_2) + \frac{\alpha_3}{T_L} \cdot \Delta F_L\end{aligned}$$

This will certainly generate consistent surfaces if the threshold for each mechanism is a constant or a function of the response of that mechanism only:

$$T_{RG} \equiv T_{RG}(F_{RG}); T_{GB} \equiv T_{GB}(F_{GB}); T_L \equiv T_L(F_L)$$

and the surface will then be defined by the equation:

$$(\alpha_1 - \beta_1) \int F_{RG} \frac{dF_{RG}}{T_{RG}^2} + (\alpha_2 - \beta_2) \int F_{GB} \frac{dF_{GB}}{T_{GB}^2} + \alpha_2 \int F_L \frac{dF_L}{T_L^2} = \text{constant}.$$

If the thresholds are not of the form assumed the existence of proper surfaces is problematic and requires fuller examination. If $\alpha_1 = \beta_1$, $\alpha_2 = \beta_2$ —which is the case for many proposed elements—everything is easier and the surfaces are specified by

$$\int F_L \frac{dF_L}{T_L^2} = \text{constant, or } \int^L \frac{1}{T_L^2} \cdot \frac{dL}{L} = \text{constant}.$$

and depend only on the luminance term in the element.

The line-element developed by Shklover (1955-1957) is of the general type we have been considering and makes the simplest possible assumptions. The thresholds for all three mechanisms, T_{RG} , T_{GB} , T_L were assumed to be constants, all the self-light constants a_1 , b_1 , a_2 , b_2 , l_0 were given equal small values, and the α and β constants in the two colour mechanisms were taken as equal. With his special choice of fundamentals—resembling most closely a Fick set—and with suitable numerical values for the threshold and other constants, Shklover was able to reproduce fairly well the simpler data on hue and saturation discrimination and to obtain semi-quantitative explanations of the effects of intensity and saturation on apparent hue. Critical testing of the element on comprehensive discrimination data such as MacAdam and Brown's was not attempted and could hardly have had much success.

A more interesting variant of the general element emerged from an analysis of the MacAdam ellipses made by LeGrand in 1949. He found that the ellipses, recomputed for two sets of fundamentals of König and Fick types respectively and plotted in a rectangular R - B diagram, had all their semi axes parallel to the coordinate axes. He found the lengths of the "blue" semiaxes depended substantially only on the blue tristimulus value. This could be regarded in the framework of the general element as meaning that the green-blue differencing mechanism had degenerated into a Helmholtz type mechanism with constant threshold

$$\xi_{GB}^2 = \frac{1}{T_{GB}^2} \left\{ \alpha_2 \frac{\Delta G}{G + a_2} - \beta_2 \frac{\Delta B}{B + b_2} \right\}^2 \rightarrow \frac{1}{T_B^2} \left\{ \frac{\Delta B}{B + b_2} \right\}^2$$

$$\text{as } \alpha_2 \rightarrow 0, \beta_2 \rightarrow 1$$

$$T_{GB} = T_B = \text{constant}.$$

As the MacAdam data referred to a plane of constant luminance, they provided no information on the luminance term in the line-element. The red-green colour-differencing mechanism must then be responsible for the variation of the "red" semi-axes of the ellipses, and LeGrand's plot of the red semi-axes against the ratio R/G is shown in Fig. 3 which refers to the case of Fick type

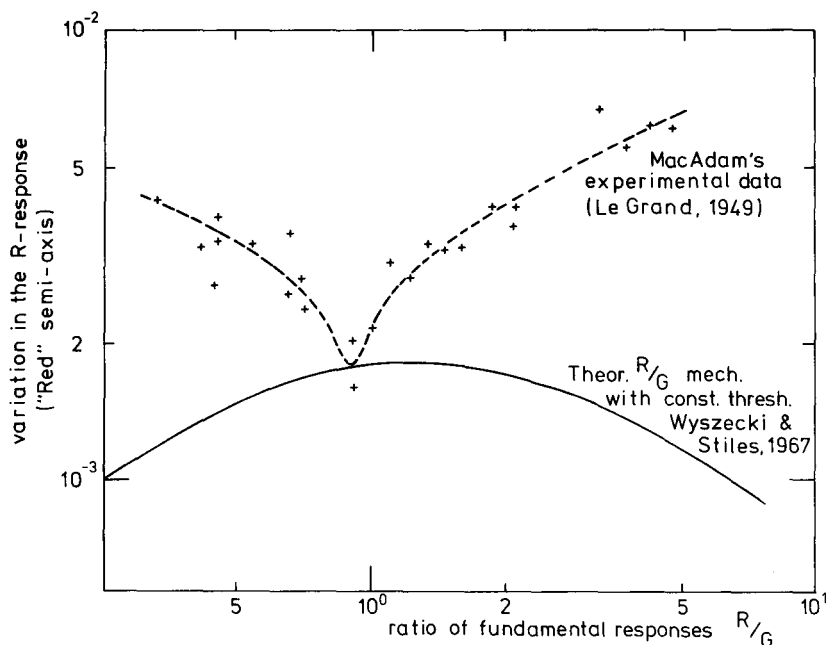


Fig. 3 Dependence of the R -semi-axis on the ratio of the fundamental responses R/G . Experimental course (MacAdam data) according LeGrand (1949). Theoretical curve expected on the assumption that the R/G mechanism has a threshold independent of R and G .

fundamentals. The semi-axis is minimal very nearly at $R=G$ and increases almost symmetrically on either side. The lower continuous curve shows the completely opposite variation which would be expected from a red-green colour-differencing mechanism of the form I showed earlier with $\alpha_1 = \beta_1$ if the threshold T_{RG} had a constant value independent of R and G . An adjustment of the Fick fundamentals towards the König type would in fact move the computed curve sufficiently to the left to bring it into symmetry with the then slightly different semi-axis curve still centred approximately at $R=G$. With these modifications it is clear that LeGrand's analysis of the MacAdam data implies for the threshold of the red-green colour-differencing mechanism, a quantity which is minimal at $R=G$ and increases in a roughly symmetrical way as R/G goes above or below unity to 3 or 4 times its minimal value. The extension of LeGrand's two term element by the addition of a third, pure luminance term

presents some problems because we do not have for the mechanism involving B , equality of α_1 and β_1 .

Considerations based on zone theory and Weber-Fechner ideas together with an analysis of Brown and MacAdam's data led Friele (1961) to propose a new line-element with interesting features. The fundamentals used were broadly of the Fick type. Extensive calculations by MacAdam (1964) to optimise the exact choice of fundamentals and constants gave improved agreement between the data and the predictions of the element. Here I want merely to comment briefly on the structure of Friele's element.

Friele's 1961 element

R-G term

$$(\Delta C_{r-g})^2 = \frac{1}{\beta^2} \left(\frac{\Delta R}{R} - \frac{\Delta G}{G} \right)^2$$

$$\beta = 0.015 \cdot \frac{R^2}{R^2 + G^2} \text{ for } R > G$$

$$= 0.015 \cdot \frac{G^2}{R^2 + G^2} \text{ for } G > R$$

$$F_{RG} = \log_e R - \log_e G \quad T_{RG} = \beta$$

Y-B term

$$(\Delta C_{y-b})^2 = \frac{1}{\gamma^2} \left(\frac{\Delta R}{2R} + \frac{\Delta G}{2G} - \frac{\Delta B}{fB} \right)^2$$

$$\gamma = 0.015 \text{ for } G < 2.5 B$$

$$= 0.015 \frac{G}{2.5 B} \text{ for } G > 2.5 B$$

$$f = 5/3$$

$$F_{YB} = \frac{1}{2} (\log_e R + \log_e G) - \frac{1}{f} \log_e B \quad T_{YB} = \gamma$$

Of the two colour-differencing terms, the red-green has close affinities with the corresponding LeGrand expression, Friele's quantity β playing the role of the threshold T_{RG} . Its minimum at $R=G$, and symmetrical increase on either side of this value resemble those found by LeGrand except that the increase is rather smaller. The yellow-blue differencing term however is more complex. The yellow response is obtained by compounding the red and green tristimulus values after the non-linear step represented by taking the logarithm, and the second member of the differencing process incorporates a factor $(1/f)$. With $f=5/3$, the yellow-blue term does not vanish when the two stimuli being discriminated differ only in luminance. This raises the questions already noted in connection with the simpler colour-differencing term when the constants α and β are not equal. The value $f=5/3$ was chosen by Friele to suit best the

Brown-MacAdam ellipsoids; for the mean observer data of Brown the value $f=1.0$ was preferred and with this choice the yellow-blue term does vanish for pure brightness discrimination. The quantity γ corresponding to the $Y-B$ threshold depends on the ratio G/B but is independent of R . In this case therefore the threshold cannot be regarded as determined by the mechanism's own response.

"Brightness" term of Friele's 1961 element

$$(\Delta L)^2 = \frac{1}{4\alpha^2} \left(\frac{\Delta R}{R} + \frac{\Delta G}{G} \right)^2$$

$$\alpha = 0.015$$

$$F_L = \log_e R + \log_e G \quad T_L = 2\alpha$$

The "brightness" term itself is striking. For the surfaces of equal small-step brightness when $f=1$, it gives the equation:

$$RG = \text{constant}$$

which is reminiscent of the corresponding results for the Helmholtz and the modified Helmholtz elements:

$$R.G.B. = \text{constant (Helmholtz)}$$

$$R^{0.61} \cdot G^{0.37} \cdot B^{0.02} = \text{constant (Stiles)}$$

Thus non-additivity of small-step brightness reappears with Friele's element. This is related to the consistent introduction of the non-linear, logarithmic step in all terms of the element. If $f=5/3$, it is not clear that there exist surfaces of equal small-step brightness.* As there are no constants playing the part of self-light constants, the element cannot be applied at very low levels.

In 1965, Friele introduced some elaborations of his line-element, and I am not sure how far my comments still apply to the new form. There is no doubt however that Friele's expressions, both the original and revised forms, have had remarkable success in organising the empirical discrimination data into an inductive line-element.

* Dr. Friele kindly sent me some observations on this point following my lecture. He adopts for his line-elements as the *definition* of surface of equal small-step brightness the condition $\Delta L=0$, where ΔL is the "brightness" term in the element. It then follows that $RG=\text{constant}$ represents the surfaces of equal small-step brightness of the 1961 elements, whatever the value of f . The definition of a small-step brightness match used throughout my lecture is the one introduced by Helmholtz and generalized by Schrödinger and, with it, the surfaces of equal small-step brightness for $f \neq 1$ —if they exist as consistent surfaces—will certainly differ from $RG=\text{constant}$. Dr. Friele observes that for his 1965 element his definition leads to surfaces specified by $R^2 + G^2 = \text{constant}$. For this element, I find that the Helmholtz-Schrödinger definition leads to the same result.

I mentioned earlier that the De Vries—Rose fluctuations concept of signal to noise ratio together with Van der Velden's suggestion of a dead-time, were used by Trabka in explaining the way the increment threshold of a Helmholtz-type mechanism increases with field intensity. In the early sixties, Bouman, Vos and Walraven (1963) used similar ideas in a fluctuation theory of luminance and chromaticity discrimination, in which the mechanisms were of the colour-differencing and luminance types. In the further development of this theory they have derived a fascinating, if complicated, line-element which embodies extensions of the ideas just mentioned, and yields predictions of the colour discrimination of dichromatic and anomalous trichromatic, as well as of normal trichromatic vision. As Drs Vos and Walraven are presenting their line-element in a paper at this Symposium, I need not attempt any summary in a few words. I would add however that this work seems to me an example, *par excellence*, of the inductive approach to the line-element.

I realise I have omitted in this lecture any reference to modifications or supplements to the line-element to cover important problems associated with different surround conditions, assessment of supra-threshold colour differences, and acceptability criteria as distinct from perceptibility thresholds. Something should perhaps be said in a historical review on the modification needed for perception in parafoveal areas of the retina, where the scotopic or rod mechanism of vision is certainly involved. Interest has turned mainly on the distortions of brightness and colour matching that occur, rather than on colour discrimination in the parafovea about which little is known. König in the eighteen-nineties had ideas on the relation between the Weber-Fechner fraction and the Purkinje shift, and more recently (Stiles, 1944) analogous ideas have been cast in line-element form. The contributions of all the cone mechanisms are lumped together in a single photopic term of Helmholtz type and added to a similar scotopic term:

$$ds^2 = \left\{ \frac{1}{\sigma} \frac{\Delta S}{S + S_0} \right\}^2 + \left\{ \frac{1}{\phi} \frac{\Delta P}{P + P_0} \right\}^2$$

The photopic and scotopic V_λ curves provide the fundamentals, and the limiting Weber-Fechner fractions—the scotopic σ equalling some thirty times the photopic ϕ —and the self-light constants S_0 and P_0 are provided by parafoveal increment threshold measurements. The Helmholtz principle of small-step brightness-matching applied to this crude line-element reproduces fairly well the way in which the experimental V_λ curve, as measured by Walters and Wright (1943), is modified in shape and shifts in the Purkinje sense as the intensity level is raised. But if we refine on this element by replacing the single photopic term by three cone terms to bring in colour perception:

$$\left(\frac{1}{\phi} \frac{\Delta P}{P + P_0} \right)^2 \rightarrow \left\{ \frac{1}{\rho} \frac{\Delta R}{R + a} \right\}^2 + \left\{ \frac{1}{\gamma} \frac{\Delta G}{G + b} \right\}^2 + \left\{ \frac{1}{\beta} \frac{\Delta B}{B + c} \right\}^2$$

we then have a four-dimensional colour-matching system in the parafovea. There is no difficulty at high brightness levels, where the large Weber-Fechner fraction of the rods makes their contribution negligible, nor at very low levels where the relatively high self-light constants of the cone mechanisms make their contributions negligible. But in an intermediate region it seems that colour matches with three colorimetric controls are still adequate in practice, although the additivity principle is no longer valid for such matches. One attempt (Stiles, 1955) to preserve the line-element was made by assuming the match to correspond to a minimal and not necessarily a zero value of ds^2 . If the controls of a trichromatic colorimeter supplying the comparison field are Q_1 , Q_2 , Q_3 the conditions for a complete match become:

$$0 = \frac{\partial}{\partial Q_1} [ds^2] = \frac{\partial}{\partial Q_2} [ds^2] = \frac{\partial}{\partial Q_3} [ds^2].$$

This procedure has had some success in explaining the effect of rod intrusion in modifying large-field colour-matching as the intensity level is reduced. One feels it would have appealed to Schrödinger to see the trichromatic principle reappearing in the form of three differential operators in place of three variables. This way of using a line-element has not, to my knowledge, been applied to Clarke's extensive data (1960, 1963) on parafoveal colour-matching. Clarke's own theoretical discussion of his results rest on the ideas of a scotopic white added to the match and on rod-cone interaction, and are less directly connected with a line-element.

Most of what I have put forward in this lecture can have no more than a mildly historical interest. Unfortunately I have not been in close touch with the developments of the last few years, but the impression I have is that, in parallel with the excellent and intensive studies of colour discrimination under conditions which are of practical importance, there is still room for more primitive investigations, under perhaps extreme and diverse experimental conditions of no particular applicational interest, which may enlarge the scope of the line-element idea.

REFERENCES

- Abney, W. de W. (1913), *Researches in colour vision and the trichromatic theory*. London, Longmans, Green.
- Alpern, M., Rushton, W.A.H., and Torii, S. (1970), Signals from cones. *J. physiol.* 207, 463-475.
- Bouma, P. J. (1946), *Kleuren en kleurenindrukken*. Amsterdam, Meulenhoff.
- Bouman, M. A., Vos, J. J., and Walraven, P. L. (1963), Fluctuation theory of luminance and chromaticity discrimination. *J. opt. soc. Amer.* 53, 121-128.
- Boynton, R. M., Ikeda, M., and Stiles, W. S. (1964), Interactions among chromatic mechanism as inferred from positive and negative increment thresholds. *Vision res.* 4, 87-117.
- Brown, W. R. J., and MacAdam, D. L. (1949), Visual sensitivities to combined chromaticity and luminance differences. *J. opt. soc. Amer.* 39, 808-834.

- Brown, W. R. J. (1952), The effect of field size and chromatic surroundings on color discrimination. *J. opt. soc. Amer.* 42, 837-844.
- Clarke, F. J. J. (1960), Extra-foveal colour metrics. *Opt. acta* 7, 355-384.
- Clarke, F. J. J. (1963), Further studies of extrafoveal colour metrics. *Opt. acta* 10, 257-284.
- De Vries, H. (1943), The quantum character of light and its bearing upon the threshold of vision, the differential sensitivity and visual acuity of the eye. *Physica* 19, 533-564.
- Friele, L. F. C. (1961), Analysis of the Brown and Brown-MacAdam colour discrimination data. *Farbe* 10, 193-224.
- Friele, L. F. C. (1965), Technical application of colour measurement: Standardization aspects. *Proc. Internat. Colour Meeting, Luzern*, 555-560.
- Gibson, K. S., and Tyndall, E. P. T. (1923), Visibility of radiant energy. *Bull. Bur. Standards* 19, 131.
- Guth, S. L., Alexander, J. V., Chumbly, J. I., Gillman, C. B., and Patterson, M. M. (1968), Factors affecting luminance additivity at threshold among normal and color-blind subjects and elaboration of a trichromatic-opponent colors theory. *Vision res.* 8, 913-928.
- Guth, S. L., Donley, N. J., and Marrocco, R. T. (1969), On luminance additivity and related topics. *Vision res.* 9, 537-575.
- Hecht, S. (1932), A quantitative formulation of colour vision. Report on a joint discussion on vision. London, *Phys. and Opt. soc.*, 126-
- Helmholtz, H. von (1891), Versuch einer erweiterten Anwendung des Fechnerschen Gesetzes im Farbensystem. *Z. Psychol. Physiol. Sinnesorgane* 2, 1-30.
- Helmholtz, H. von (1892), Versuch, das psychophysische Gesetz auf die Farbenunterschiede trichromatischer Auge anzuwenden. *Z. Psychol. Physiol. Sinneorgane* 3, 1-20.
- Helmholtz, H. von (1896), *Handbuch der physiologischen Optik*. 2. umgearb. Aufl. Hamburg, etc., Voss.
- Hunter, R. S. (1942), Photoelectric tristimulus colorimetry with three filters. *J. opt. soc. Amer.* 32, 509-538.
- Judd, D. B. (1935), A Maxwell triangle yielding uniform chromaticity scales. *J. opt. soc. Amer.* 25, 24-35.
- König, A., und Dieterici, C. (1884), Zur Kenntnis dichromatischer Farbensysteme. *Ann. Phys.* 22, 567-578 *Gräf. Arch. Ophthalmol.* 30, 155-170.
- König, A., und Brodhun, E. (1889), Experimentelle Untersuchungen über die psychophysische Fundamentalförmel in Bezug auf den Gesichtssinn. 2. Mittlg S.B. Preuss. Akad. Wissenschaften, 641-644.
- LeGrand, Y. (1949), Les seuils différentiels de couleurs dans la théorie de Young. *Rev. opt.* 28, 261-278.
- LeGrand, Y. (1948), *Lumière et couleurs*. Paris, Ed. "Revue d'Optique".
- MacAdam, D. L. (1942), Visual sensitivities to color differences in daylight. *J. opt. soc. Amer.* 32, 247-274.
- MacAdam, D. L. (1944), On the geometry of colour space. *J. Franklin Inst.* 238, 195-210.
- MacAdam, D. L. (1964), Analytical approximations for color metric coefficients. II. Friele approximation. *J. opt. soc. Amer.* 54, 247-256.
- Nickerson, D., and Newhall, S. M. (1943), A psychological color solid. *J. opt. soc. Amer.* 33, 419-421.
- Rose, A. (1948), The sensitivity performance of the human eye on an absolute scale. *J. opt. soc. Amer.* 38, 196-208.
- Saunderson, J. L., and Milner, B. I. (1946), Modified chromatic value color space. *J. opt. soc. Amer.* 36, 36-42.
- Schrödinger, E. (1920), Grundlinien einer Theorie der Farbmétrie im Tagessehen. *Ann. Physik* 63, 397-427; 427-456; 481-520.
- Scofield, F. (1943), A method for determination of color differences. *Nat. Paint Varnish Lacquer Assoc., Science circ. no.* 664.
- Shklover, D. A. (1955), Méthodes photoélectriques et appareils pour mesurer la couleur. *Proc. CIE Zürich*. 18 pp.
- Shklover, D. A. (1957), The equicontrast colorimetric system. *Proc. Symposium 'Visual*

- problems of colour', Teddington, 605-614.
- Silberstein, L. (1943), Investigation on the intrinsic properties of the color domain. *J. opt. soc. Amer.* 33, 1-10.
- Silberstein, L., and MacAdam, D. L. (1945), The distribution of color matchings around a color center. *J. opt. soc. Amer.* 35, 32-39.
- Stiles, W. S. (1944), Current problems of visual research. *Proc. phys. soc.* 56, 329-356.
- Stiles, W. S. (1946), A modified Helmholtz line-element in brightness-color space. *Proc. phys. soc.* 58, 41-65.
- Stiles, W. S. (1955), Remarks on the line-element. *Farbe* 4, 275-279.
- Stiles, W. S. (1967), Newton lecture: "Mechanism concepts in colour theory". *J. Colour Group* nr 11: 106-123.
- Trabka, E. A. (1968), On Stiles' line element in brightness-color space and the color power of the blue. *Vision res.* 8, 113-134.
- Trabka, E. A. (1969), Effect of scaling optic-nerve impulses on increment thresholds. *J. opt. soc. Amer.* 59, 345-349.
- Van der Velden, H. A. (1949), Quanteuze verschijnselen bij het zien. *Ned. t. natuurk.* 15, 145-150.
- Vos, J. J., and Walraven, P. L. (1971), A zone-fluctuation line-element describing colour discrimination. *Proc. Symp. on 'Color metrics', Driebergen.*
- Walters, H. V., and Wright, W. D. (1943), The spectral sensitivity of the fovea and extra-fovea in the Purkinje range. *Proc. roy. soc. B* 131, 340.
- Wright, W. D. (1941), The sensitivity of the eye to small colour differences. *Proc. phys. soc.* 53, 93-112.
- Wyszecki, G., and Stiles, W. S. (1967), *Color science*. New York, Wiley.

DISCUSSION

Due to lack of time, there was no discussion after this paper. The theme was picked up again, however, in the general discussions.

2 COLOR DIFFERENCES ASSESSED BY THE MINIMALLY-DISTINCT BORDER METHOD

R. M. BOYNTON and H. G. WAGNER

*Center for Visual Science, University of Rochester, Rochester,
New York, U.S.A.*

INTRODUCTION

The purpose of this research is (a) to investigate a new method for evaluating the sensory difference between various colors; and (b) to determine whether the results of the experiment can be represented satisfactorily in a 2-dimensional space.

Previous attempts to measure the sensory differences between colors have been of two general kinds. In the first, pioneered by MacAdam (1942) and extended by Brown (1957), just-noticeable differences were established in chromaticity space in the neighborhood of various samples, with the results being expressed by ellipses whose size and orientation were descriptive of these differences for various directions in chromaticity space. This is a satisfactory method for assessing small differences, but is difficult to apply without ambiguity to the assessment of large ones. Most studies of large differences have relied upon direct estimation techniques, where for example the subject may be asked to judge the ratio of the perceived difference between two colors, as compared to some fixed reference. The most recent examples of this sort are studies by Ramsey (1968) and Burnham, Onley, and Witzel (1970). The main difficulty with the general application of such procedures is well stated by Burnham *et al.*:

"However, in actually making such judgments, there are limits beyond which a truly random selection [of stimuli] cannot reasonably be used. It became clear in preliminary observations, for example, that judgments of the visual interval between qualitatively different hues, such as red and green, were difficult if not impossible. Any randomization of color pairs should probably be confined to relatively small regions of the three-dimensional color domain . . ." (p. 1410).

The perception of a contour is generally mediated by a luminance difference over space which occurs with sufficient suddenness to define a border. Although

the addition of chromatic differences does not usually add much to the number or strength of contours that are perceived, there are special cases where chromatic differences alone may be shown to support the perception of a border.

If two fields of equal luminance, having identical spectral distributions, are precisely juxtaposed, there will be no physical discontinuity of radiance at their junction. Therefore no detector, whether physical or biological could determine the presence of a contour. In the more general case where a univariantly responding detector is used, having a fixed spectral sensitivity, (e.g., a photocell), many physically different stimuli could produce the same response of the detector to each part of the field, and no border would then be discerned. This would be true also for scotopic vision, for any two precisely juxtaposed stimuli for which $\int L_{e1\lambda} V_{\lambda}' d\lambda = \int L_{e2\lambda} V_{\lambda}' d\lambda$, where $L_{e1\lambda}$ is the radiance of one field, V_{λ}' is the scotopic luminous efficiency function, and the terms on the right-hand side of the equation are similarly defined. The observer is assumed to have a scotopic visual sensitivity exactly matching that of the standard observer.

In photopic vision, where three classes of photoreceptors are involved, it is usually impossible to produce the same degree of activation of all three systems with two physically different stimuli. The contour perceived between them will never disappear, regardless of their relative radiances, provided that photopic radiance levels are employed, and metameric pairs (or ones very nearly so) are avoided.

In our work, we focus attention upon the minimally-distinct contour that can be perceived between two such colors. This occurs for some particular radiance value of the one stimulus which is set relative to the other. We then ask: how distinct is the residual contour at this minimally-distinct setting? We evaluate this by a matching method. The results are then analyzed by analysis of proximities, a non-metric technique of multidimensional scaling developed by Shepard (1962, 1966) and Kruskal (1964a, b).

PROCEDURE

The apparatus allows the projection, upon a calcium carbonate block, of six separate fields of light which can be independently adjusted for spatial position. The fields to be assessed are at the bottom of the four-part field shown in Fig. 1. For example, suppose that 490 nm on the left is to be compared with 580 nm on the right. In this case, fields 3*W* and 4*W*, capable of supplying white light, are extinguished. The wavelength of field 3*λ* is set to 490 nm, that of 4*λ* to 580 nm. Field 3*W* is adjusted for luminance to be approximately 80 *td*, and the two fields are adjusted to provide the best possible juxtaposition at the vertical dividing line. The subject then adjusts field 4*λ* to produce a minimally-distinct border between it and the field on the left.

The upper fields are used to assess the distinctness of the contour formed at

the bottom. With field $2W$ set at 80 td of white light, the subject adjusts field $1W$, also white and of the same chromaticity, to form an achromatic border which matches the lower, chromatic one for distinctness. This can be done two ways: by making the luminance of $1W$ higher (positive contrast) or lower (negative contrast) than field $2W$. The contrast C at the border in the upper field is defined as

$$C = (L_1 - L_2) / (L_1 + L_2),$$

where L_1 is the luminance of field 1, L_2 is the luminance of field 2. C can vary from 0 when $L_1 = L_2$ to 1.00 (when L_1 or L_2 is zero). Both positive and negative contrasts were used, once each with the stimuli below in both possible left-right positions. No significant differences were obtained as a function of left-right positions; but there were some differences in the absolute values for positive vs. negative contrast settings. Nevertheless, the results for the two conditions of contrast were combined and the data to be analyzed by the Shepard-Kruskal technique were therefore averages of the values obtained for all four conditions (the two methods of contrast setting, for each left-right position). The contrast will be expressed here as a percent multiplying C by 100.

In order to minimize the chromatic aberration of the eye, an achromatizing lens (Wysecki and Stiles, 1967, p. 212) was used; without this, many of the comparisons would have been impossible. The horizontal black line dividing the upper and lower fields was deliberately introduced to minimize induction effects between the upper and lower parts of the stimulus configuration.

A total of 17 stimuli were used. There were monochromatic stimuli at

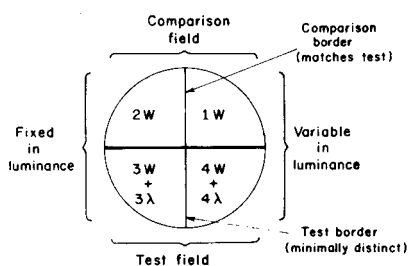


Fig. 1. Configuration of the stimulus fields used in the experiment. The distinctness of the chromatically-supported border at the bottom is evaluated with an achromatic border, at the top, which is set to appear equally as distinct.

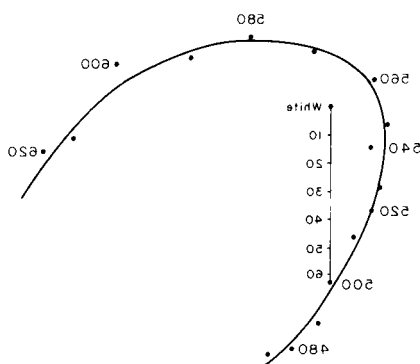


Fig. 2. Results of the Shepard-Kruskal analysis of proximities. The non-labelled points are, in each case, the appropriately intervening ones. In the original figure, the distance D between the White point and 500 nm was 63 mm.

10 nm steps from 470 to 620 nm inclusive, plus White ($x=0.29$, $y=0.32$). This produced a matrix of $(17 \times 16)/2 = 136$ values for contrast in the upper field as the primary data of the experiment.

Data reported here are for one subject (*GW*): at the time this manuscript is being prepared, data are being obtained on four more subjects.

RESULTS

Is it possible to arrange 17 points in a two-dimensional space, so that the distances between all possible pairs of points is related to the contrast values of the experiment? To find out, the Shepard-Kruskal scaling method was employed, yielding the result shown in Fig. 2. The values of stress for dimensions ranging from 1 to 4 are 0.398, 0.277, 0.262, and 0.254, meaning that two dimensions are necessary and sufficient for representing the data. Although the stress values seem rather high, the very orderly arrangement of the spectral stimuli, and the seemingly reasonable location of White, is encouraging and prompted further analysis. The distance between the White point and the various spectral stimuli was computed, and is shown as a function of wavelength in Fig. 3. This figure clearly resembles saturation discrimination functions obtained by other methods. The measured contrast for these stimulus pairs is also shown on the same figure, and it appears that measured contrast and distance in the figure are correlated, although probably non-linearly.

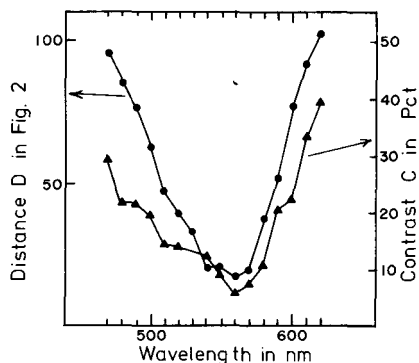


Fig. 3. Distance in Fig. 2, between the White point and all spectral wavelengths, is plotted as a function of wavelength. Also shown is the matching contrast of the upper fields which produced a border of matching distinctness.

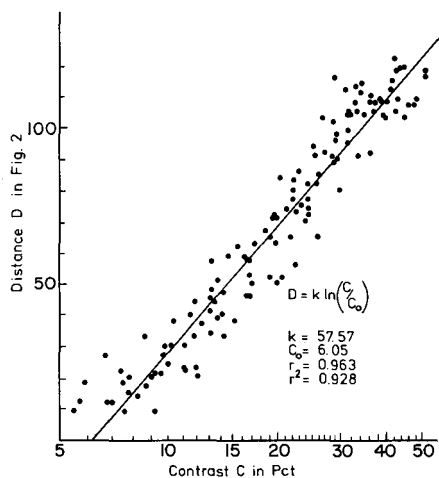


Fig. 4. Scattergram for all 136 pairs of stimuli, showing the relation between distance in the diagram of Fig. 2 and the contrast between the matching upper fields of Fig. 1, plotted with logarithmic spacing.

This correlation was computed for all 136 pairs of values of D and C and was found to be 0.939 (Pearson product-moment), meaning that distance on the diagram accounts for $r^2=88.2$ pct of the experimental variance. The relation may be improved by taking the logarithm of the contrast values. The scattergram for D vs. $\log C$ is shown in Fig. 4, where it will be seen that more than 96 pct of the experimental variance is accounted for.

Results from additional subjects. The experiment just described has been repeated in its entirety, using two additional subjects. These additional data are given in Tables I, II, and III, where they are compared with the results for GW , whose data are shown in Figs 2, 3, and 4.

Adding additional dimensions to the analysis necessarily reduces stress and improves the correlation between D and $\ln C$. We judge, from Table I, that the improvement obtained in going from two to three dimensions is not sufficient to justify using the additional dimension. Table I also shows that the data of RMB are nearly as good as those of GW , but show slightly higher values of stress for the two-dimensional analysis, and slightly lower correlations between D and $\ln C$. Subject FW yields poorer data, but his Shepard-Kruskal diagram closely resembles the other two subjects.

Table II compares the results of the 17-stimuli analysis with an analysis using only 9 of the 17 stimuli (470-610 nm, in 20 nm steps, plus white). Stress values are lower, as is characteristic of this index when the number of points is reduced, but the correlations are quite similar, as are the resulting Shepard-Kruskal diagrams (not shown). Thus the result is not critically dependent upon the number of stimuli chosen for analysis.

Table III provides another way to compare data among subjects, and again shows that the individual differences are not large.

Table I. Influence of Number of Dimensions

Subj.	Number of Dimensions Used in the Analysis	Value of Stress Obtained (see Indow and Ohsumi, 1971, for Equ.)	Product-Moment Coefficient of Correlations Between Diagram Distance (D) & $\ln C$
RMB	1	0.118	.915
	2	0.074	.956
	3	0.059	.961
FW	1	0.220	.856
	2	0.126	.912
	3	0.087	.933
GW	1	0.158	.907
	2	0.060	.976
	3	0.046	.980

Table II. Influence of Number of Points

Subj.	17		9	
	Value of Stress	Correlation	Value of Stress	Correlation
RMB	0.074	.956	0.025	.957
FW	0.126	.912	0.094	.871
GW	0.060	.976	0.020	.967

Table III. Correlation Between In Contrast and Distance (D) in the Diagram Describing Results of Same and Different Subjects

		<i>Distance (D)</i>			
In Contrast	RMB	.956	.912	.918	.942
	FW	.884	.912	.883	.903
	GW	.946	.940	.976	.971

DISCUSSION

The results of this experiment suggest that there is a relation between the distinctness of a minimally-distinct border, and the sensory difference between the stimuli forming that border. The evidence lies in the similarity of the diagram of Fig. 2 to psychological color diagrams obtained by other methods. It must be emphasized, however, that sensory differences are not being directly assessed, in that subjects are asked only to judge the distinctness of a border separating two fields of differing chromatically, without making a judgment about how different two fields appear. Our ability to characterize these differences with such high precision in a two-dimensional representation suggests that the minimum-border method somehow taps the basic two-dimensional character of chromatic variations where intensity is controlled.

In terms of the model of brain function introduced by Boynton and Kaiser (1968) and expanded upon by Kaiser, Boynton, and Herzberg (1971), the results of this experiment have implications, with respect to interactions among hypothesized chromatic brain elements of different types, concerning how these combine to give rise to chromatically-determined contour.

REFERENCES

- Boynton, R. M., and Kaiser, P. K. (1968), Vision: the additivity law made to work for heterochromatic photometry with bipartite fields. *Science* 161, 366-368.
- Brown, W. R. J. (1957), Color discrimination of twelve observers. *J. Opt. Soc. Amer.* 47, 137-143.
- Burnham, R. W., Onley, J. W., and Witzel, R. F. (1970), Exploratory investigation of perceptual color scaling. *J. Opt. Soc. Amer.* 60, 1410-1420.
- Indow, T., and Ohsumi, K. (1971), Multidimensional mapping of sixty Munsell colors by non-metric procedure. *Proc. Symp. on "Color-metrics"*, Driebergen
- Kaiser, P. K., Herzberg, P. A. and Boynton, R. M. (1971), Chromatic border distinctness and its relation to saturation. *Vision Res.* 11, 953-968
- Kruskal, J. B. (1964a), Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29, 1-27.
- Kruskal, J. B. (1964b), Nonmetric multidimensional scaling: a numerical method. *Psychometrika* 29, 28-42.
- MacAdam, D. L. (1942), Visual sensitivities to color differences in daylight. *J. Opt. Soc. Amer.* 32, 247-274.
- Ramsay, J. O. (1968), Economical method of analyzing perceived color differences. *J. Opt. Soc. Amer.* 58, 19-22.

2 DISCUSSION

- Shepard, R. N. (1962), The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika* 27, 125-139 and 219-246.
- Shepard, R. N. (1966), Metric structure in ordinal data. *J. Math. Psychol.* 3, 287-315.
- Wyszecki, G., and Stiles, W. S. (1967), *Color Science*. New York, Wiley.

DISCUSSION

Guth: If it is true that in your experiment a physiological achromatic system is isolated, then would not you expect distances between the points along your spectrum locus to resemble wavelength discrimination for an equal *CIE* luminance (as opposed to brightness) spectrum?

Boynton: I would not say that the minimally distinct border technique isolates the achromatic system. In these experiments, after the subject makes an initial setting (which reduces the border to a minimum), the border that remains is presumably due entirely to chromatic differences.

Guth: That is what I meant.

Boynton: The spacing of our stimuli was a bit too coarse to justify the type of analysis you suggest. However, Fig. 2 shows that, roughly speaking, there is such a correlation. For example, in the yellow part of the spectrum, where sensation varies rapidly with wavelength, the distances between the wavelengths are in fact greater than elsewhere. Analyses that we have done which are specifically concerned with saturation, show a very good relation between both sets of data. In Fig. 3, for example, the distance between the white point and the spectral points is clearly correlated with saturation discrimination functions. Our method, I think, provides a new way of assessing saturation which gets around some of the criterion and definitional problems that certain other methods tend to raise.

Kowaliski: By which specific experimental technique have you avoided the effects of local adaptation which tends to make disappear the separation line on the one hand, or to make it float and dissociate on the other hand in two strips of complementary colors, specially with magentas, but less pronounced with all other colors.

Boynton: For low contrasts especially, the contour tends to fade due to the Troxler effect (which is related to the stabilised image phenomenon) with prolonged fixation. In some experiments we have interrupted the stimulus for about a quarter of a second, every four seconds, which seems to help. Personally, I prefer deliberately to move my fixation up and down, which accomplishes very much the same objective. One of the advantages of using a matching technique as we did, where the observer evaluates the apparent distinctness of the lower border by making a match between it and the upper one, is that the upper and lower contours tend to fade in very much the same way. In fact I would predict that, if you did a stabilized image type of experiment, the fade-out time would be very close for chromatically and achromatically supported contours which match for distinctness in our experimental situation. Also,

from time to time within a session, and from one day to the next, the apparent subjective vividness of a given objective border will vary depending on the state of the subject, and this also tends to vary in much the same way between the upper and lower field. Thus we believe the matching technique tends to get rid of the experimental variance which might otherwise be introduced by that problem. The next thing we are going to do is use an absolute judgment technique. We will use a 7-points scale of distinctness, a variation of a Stevens type method. The subjects will make absolute judgements of apparent border strength on such a scale, for all of the combinations of stimulus pairs that we have used before. We'll subject these data to the same analysis, and I suspect that they will produce pretty much the same diagram. But I expect the data will be somewhat more variable, for the reasons mentioned.*

Kowaliski: How long was your observation time?

Boynton: This is up to the subject. He is adjusting the luminance of the upper right hand fields as he looks at the field, so the viewing time is variable. Typically this takes 5 or 6 seconds. The lengths of the experimental sessions also vary a good deal. Some of us can take it for longer than others. Thus the viewing time was not precisely standardized. Possibly it should be in extensions of this sort of work.

Indow: I have made a similar attempt mentioned in the second comment of Prof. Guth with the plotting reported by Shepard and Kruskal on the Ekman data. The result was in fairly good agreement with the ordinal wavelength discrimination data (Indow and Takagi, 1968).

Boynton: You will recall, if you've read Prof. Indow's abstract, that he also uses the Kruskal technique in his work, and if time permits later in the symposium, it might be possible to go more into details about the method.

Jaeckel: Variability should increase if you change from using the eye essentially as a null-point detector, in your present method, to a scale indicator, at which it is less efficient and consistent.

Boynton: Very well put, and I agree completely.

Clarke: Would not Dr. Boynton agree that he would expect tritan contrasts (those that a tritanope cannot see e.g. yellow vs violet) would be more affected by fixation, fading and self-adapting effects than other types of contrast i.e. red vs green or brightness differences?

Boynton: Our results cannot be generalized very far beyond the experimental conditions that we used, particularly with respect to field size, although they would probably change in somewhat predictable ways if we made the test field larger.

* In autumn, 1971, we did this experiment with one subject, RMB. The results are, indeed, very similar to those obtained using the matching technique. To our surprise, they do not seem to be more variable.

2 DISCUSSION

Vos: Having established the relation between D and log contrast, is the Kruskal non-metric analysis not quite a detour, in retrospect? Did you apply metric analysis as well, in terms of log C ?

Boynton: I think that a metric analysis is implied by the computation of the correlations that I presented to you. We've taken the distances in the diagram, a Euclidean space, and we've measured them. A great strength of Kruskal's analysis (which Shepard originally worked on) is that, despite the fact that the method discards the metric values (if you happen to have them) at the beginning of the analysis, it proves capable of restoring metric differences in the end.

The way Shepard demonstrated this was to draw a circle, upon which he spotted some points irregularly. He then measured all the distances between possible pairs of points, and put them into the analysis. The first thing the analysis does is to get rid of the metric distances and to rank-order the scores. Yet in the end you come back with something which is very close to a circle. The reason for this is that you have many more combinations of pairs of points than you minimally need to determine a circle, and there simply is no other way to satisfactorily arrange such points in a two-dimensional space (so that the rank-ordering of the distances between all possible pairs is correct) without in fact coming back again to something which is very close to a circle.

The nonmetric first step is an extremely important point for psychological data involving scaling judgements of various kinds, because often we know so little about the metric properties, if any, of the particular psychological scale on which we ask subjects to make judgements.

Vos: But here you found a metric scale: the log C scale. So why then not try a metric analysis using log contrast as your distance parameter?

Boynton: Well, that is essentially what we end with—that is what the correlations tell us. The correlations that I reported to you are Pearson product-moment correlations between metric distances in a two dimensional diagram, and the logarithm of contrast as obtained experimentally.

Guth: Can your method be used at lower intensities or to assess smaller chromatic differences?

Boynton: The intensities that we have used have varied from about 8 to 80 td —not exactly dazzlingly bright. If we were to go much lower, under our conditions of viewing, the smallest differences (e.g., between white and 570 nm) would probably go below threshold. This could be compensated to some extent by increasing area. The method can, in principle, be used to assess smaller chromatic differences; to do so would probably require the use of higher intensities and larger areas. The ultimate limitation may be related to a phenomenon well illustrated by some of the ceramic tile samples of Mr. Malkin, on display here. When the contrast (chromatic or achromatic) between two fields is very small, and their juxtaposition very precise, discriminability is

worse than if a fine dividing line is introduced between them (we have noticed this in our experiments also).

But to introduce such a dividing line would totally invalidate our method.

REFERENCES

- Indow, T., and Takagi, C. (1968), Hue-discrimination thresholds and hue-coefficients. *Jap. psychol. res.* 10, 179-190.

3 A NEW METHOD OF LARGE FIELD COLOUR MATCHING LEADING TO A MORE ADDITIVE METRIC

P. W. TREZONA

*National Physical Laboratory,
Teddington, England*

INTRODUCTION

Any assessment of the chromaticity corresponding to a given spectral power distribution by a non-visual method, requires the use of a set of colour matching functions (*cmfs*) which have themselves been determined visually to represent an average observer. Tables can be used for computation or direct instrumental measurements can be made with 3 photocell-filter combinations, the responses of which have been fitted as closely as possible to the three *cmfs*. *Cmfs* have been standardised by the *CIE*, in 1931 to represent the 2° field and in 1963 for the 10° field. Measurements on a 2° field suffer from the inherent disadvantage that the field is too small to be representative of many practical viewing situations while measurements with the 10° field encounter difficulties of additivity failure caused by rods in all fields subtending more than 2°.

Stiles (1955) demonstrated how, in a 10° field, a match can greatly change merely by altering the luminance level; this is because the extent of rod intrusion greatly increases with diminution of level. Non-additivity was shown for large and small fields off the fovea by Clarke (1960, 1963). To avoid rod intrusion the investigations of Stiles & Burch (1959) and of Speranskaya (1959), on which the 1963 standards were based, were conducted at as high luminance levels as possible; even so, certain other precautions had to be taken.

Since the 1963 large field *cmfs* are applicable only at high luminance levels there is no system suitable for large field colour matching at all luminance levels. We need a system which allows for rod intrusion by providing an extra matching stimulus for equating rod activity as well as that of the 3 chromatic cone processes. This is necessary, not because trichromatic matches are impossible or even difficult, but because they are distorted matches leading to non-additivity. Four-dimensional colour matching was first reported by

Bongard et al (1957). Apparently by trial and error, they obtained colour matches stable to changes in luminance and adaptation. It is the purpose of this paper to describe a systematic method of obtaining a 4 stimulus colour match in principle, in technique, and in some of its properties.

METHOD

Principle. The systematic method of achieving the tetrachromatic match, first described by Trezona (1970) balances rods as well as the 3 cone processes. The former takes place at a luminance level above rod but below cone threshold, obtained by viewing the whole field through a rotating sector disc run at above the critical fusion frequency, which attenuates the field neutrally. One matching stimulus is adjusted at the low level for an achromatic luminance match and the other 3 for the full colour match at the experimental level. The low-level matching stimulus (*C*) is a cyan chosen for maximum rod/cone sensitivity while *R*, *Y*, *B* are chosen for minimum rod/cone activity and to cover the colour gamut. The low and the experimental level matches are made alternately; each disturbs the previous one, but by successive approximations a match acceptable at both levels can be achieved.

Procedure. Measurements were made on the *N.P.L.* Trichromator (Stiles, 1955) which provides a bipartite field, one half, the test field, containing the test stimulus and the other, the mixture field, containing the matching stimuli—a red (*R*), green (*G*) and blue (*B*) in Stiles' original set-up. *R*, *G* and *B* can be made negative by combining them with the test stimulus, when they are called desaturating stimuli *DR*, *DG* and *DB* respectively. Modifications (Clarke and Trezona—to be published) provide a set of 4 matching stimuli in both mixture and desaturating fields, all of which can be attenuated independently. Wavelengths used are 644 nm (*R* and *DR*), 588 nm (*Y* and *DY*), 509 nm (*C* and *DC*), and 468 nm (*B* and *DB*). Values of V'_λ at these wavelengths can be seen in Fig. 1. An important feature is the sector disc, mounted between the final objective and the exit pupil. If attenuation for the "low level" match is to be achieved only by the sector and if this level must be below cone threshold, then the maximum experimental luminance level is governed by how small the sector angle can be made. On the disc used, it could be set at 0.1° , giving a reduction ratio of 3600 : 1. As it is not possible to mount the disc at the exit pupil, the profile of the sector edges and their being truly radial become important to avoid attenuating one part of the field more than another, this becoming very critical for such small angular openings. Bearing these points in mind, the edges were made of ground steel and adjusted to lie along radii to an accuracy of 5 μm . Being asymmetrical, the disc required careful balancing. Fine engravings at intervals of $\frac{1}{2}^\circ$ up to 30° allowed interpolations to 0.1° . Adjustment was by a rack and pinion mechanism and then a clamp was applied. The angle

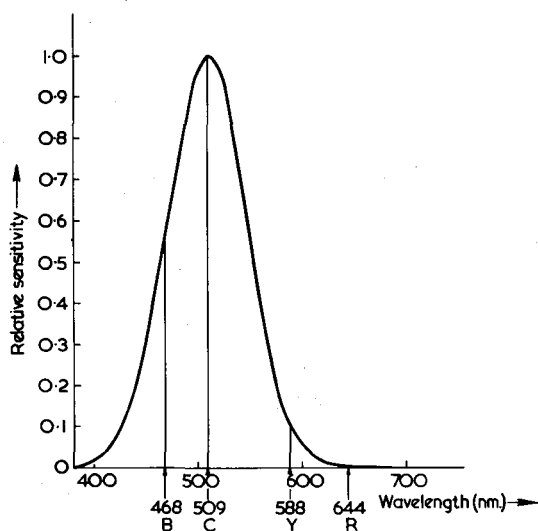


Fig. 1. Scotopic sensitivity curve V'_λ showing the relative rod sensitivity of the matching stimuli.

30° allowed the field to be seen without vignetting when the disc was stationary.

Energy measurements could be made directly by a photomultiplier (Crawford 1965), the relative spectral responsivity of which had been previously measured. However, for greater convenience, readings of scales attached to the wedges were made during the observing session, enabling all photomultiplier readings to be made together at the end of the session. A lamp of known luminous intensity can be set up at the exit pupil position to give an absolute calibration in trolands and to act as a day-to-day standard for the photomultiplier.

The two observers were normal but untrained; a short training session in trichromatic and tetrachromatic matching was given.

The "low" luminance level, although not critical, must be chosen with care. In determining the V'_λ curve for a central 20° field, Crawford (1949) found that too low a level caused great fatigue; at the lowest acceptable level, where measurement took place, the field was colourless except for long wavelengths where it was reddish (Crawford, 1962). In the present investigation, precision was lost when the "low level" was too low; matching difficulty as well as unstable matches, varying greatly from day to day, occurred when it was sufficiently high to be coloured. The criterion adopted was to have the field as bright as possible without seeing colour, except sometimes a reddish tinge. Once the observer had decided on the sector angle, this was re-set on subsequent occasions by the colleague who acted as recorder.

Each matching session followed 10 minutes of dark adaptation. Positioning of the eye pupil was set by a telescope as described by Stiles (1955).

The observer was first asked to make a trichromatic match by excluding one or other of the 4 matching stimuli. Then the sector was set on the approx-

priate small angle, the motor switched on and the observer asked to adjust *C* (or the appropriate desaturating stimulus *DC*) only. When the motor was switched off and the sector opened up, he made a trichromatic match with *R*, *Y*, *B* (or appropriate *DR*, *DY* or *DB*). Then the low level match was repeated and so on. After a few iterations, the observer started to regard the previous match as acceptable, and the recorder could see that systematic changes had ceased. At this point, the recorder started to upset the matches, moving the 3 wedges before the trichromatic match and the cyan wedge before the luminance match. Usually 5 more pairs of matches were then made.

It has been stated that the tetrachromatic match is approached from an initial trichromatic match, obtained by omitting one or other of the 4 matching stimuli. With the same test colour, a different initial matching stimulus was excluded on each of 4 different occasions, so that the tetrachromatic match was approached from the 4 extreme conditions for every test stimulus used. This is a fairly severe test of the convergence to a final match, and a test too of how rapidly it occurs. To test the generality of the method, the test colour was taken in turn from each of the 5 sections of the spectrum which occur with 4 matching stimuli. Desaturated colours are represented by a similar investigation using a continuous-spectrum white test stimulus.

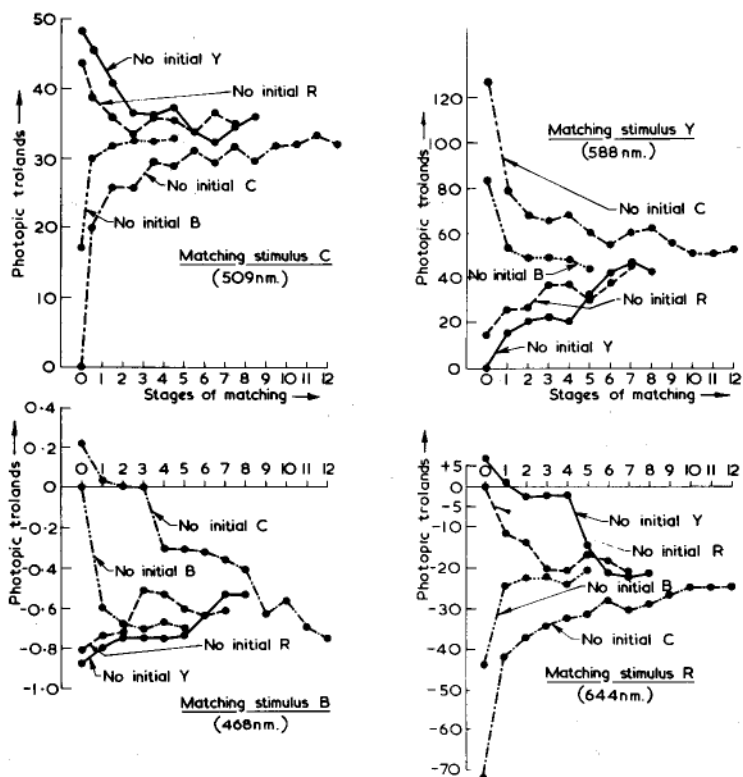
The tetrachromatic match was designed to abolish colour match change with variation of luminance level. To assess to what extent this was achieved, the technique was tested by comparing tetrachromatic and trichromatic match changes for the same test stimulus, which was first 530 nm and then 546 nm.

RESULTS

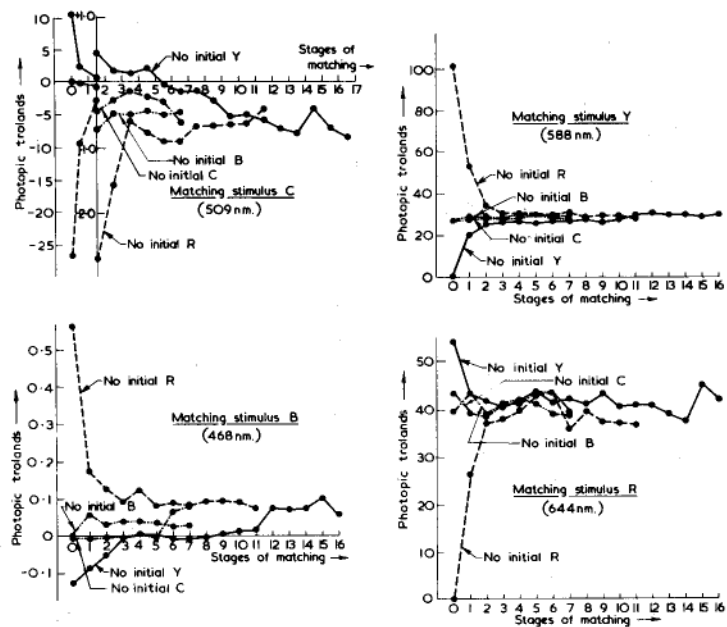
Figs 2, *a*, *b*, *c* show the results of tetrachromatic matches on test colours 530 nm, 610 nm and 485 nm respectively, performed by *PGM* in that order; he had had very little experience before the first. A second observer *ERF* obtained results very similar to those of Fig. 2*a*. In each figure there are 4 parts, one for each matching stimulus; in each part there are 4 curves, corresponding to the 4 different initial trichromatic matches and identified by the matching stimulus *missing* in this initial match. Ordinates are in photopic trolands, positive for a mixture colour and negative for a desaturating stimulus. The abscissa shows the stages of matching. Zero is the initial trichromatic match while 1, 2, 3, etc are the subsequent trichromatic matches which use *R*, *Y*, *B*; low level matches with *C* are at $\frac{1}{2}$, $1\frac{1}{2}$, $2\frac{1}{2}$, etc.

It will be seen that convergence towards the final tetrachromatic match occurs in every case, even when a matching stimulus changes from one side of the field to the other. A procedure has been adopted to facilitate this change-over. For instance, for "No initial *Y*" in Fig. 2*a*, where *R* has to be changed to *DR*, a small fixed amount of *DR* is provided before the change takes place.

2a



2b



2c

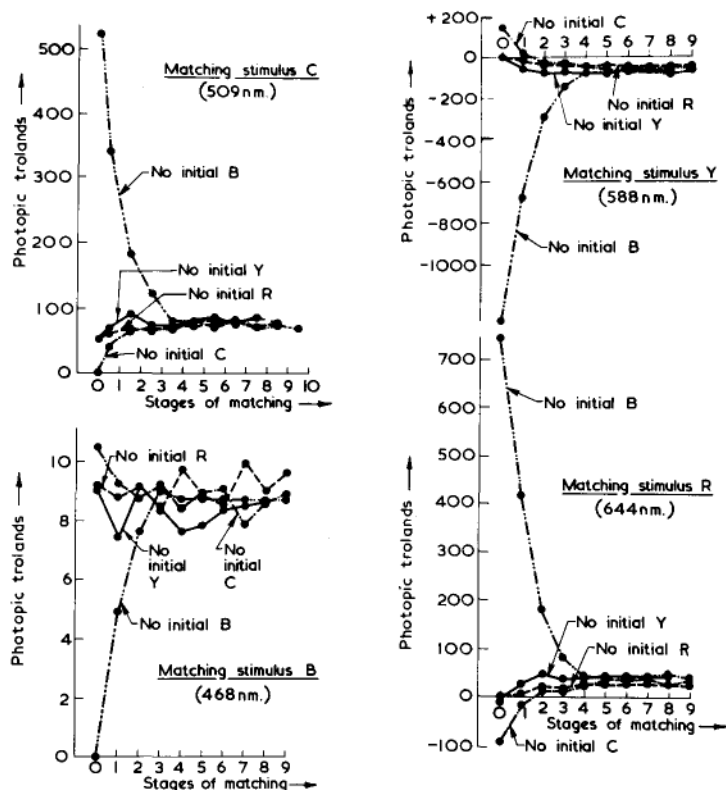


Fig. 2. Approach to the tetrachromatic match from the four extreme starting conditions. Amount of each matching stimulus required for a match on the test stimulus is plotted against the number of steps taken towards the final match.

- a. test stimulus 530 nm, retinal illumination 60 photopic trolands,
 b. test stimulus 610 nm, retinal illumination 65 photopic trolands,
 c. test stimulus 485 nm, retinal illumination 55 photopic trolands.

Then as soon as the observer sets *R* so that its energy is less than that of *DR* (equal energies previously having been determined in scale readings), he is given the *DR* control instead.

Practice helped the change-over. In Fig. 2a, the two cases involving such a cross-over, i.e. "No initial *Y*" and "No initial *C*", required many more iterations than the other two cases. But in Fig. 2c, even a double cross-over was achieved by stage 2 for "No initial *C*". The observer did this by making many successive changes in *R*, *Y* and *B*, all for a single stage of matching.

Difficulty was experienced in the violet region where, on some occasions, the observer first converged towards and then diverged from the tetrachromatic match. This was successfully dealt with and convergence established by allowing the observer only to glance at the field to avoid adapting to it and by giving

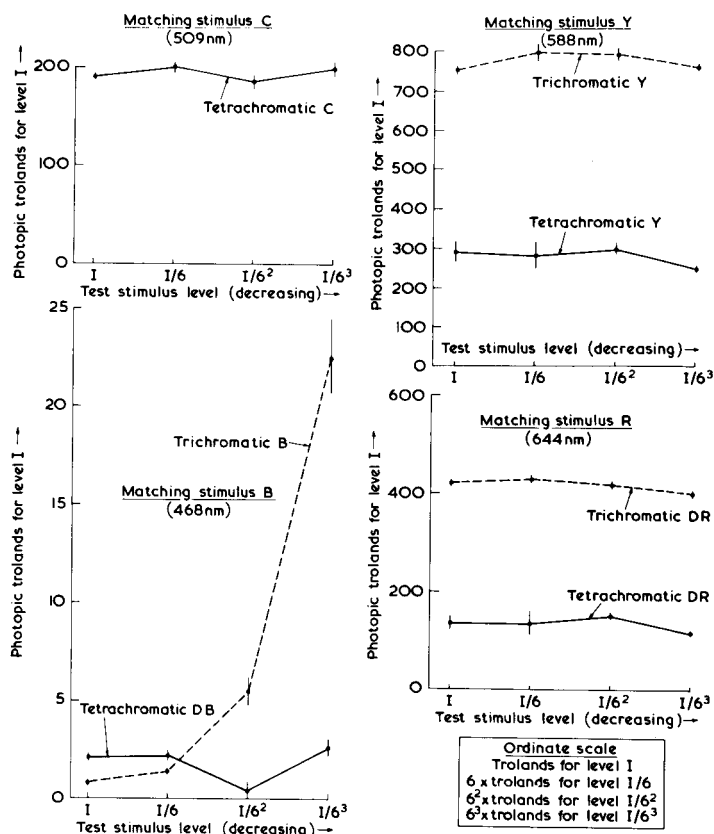


Fig. 3. Trichromatic and tetrachromatic colour matches at four levels. Test stimulus 530 nm, 400 photopic trolands at experimental level I. Lines show $\pm 2 \times$ standard error of mean.

him more frequent rests than usual. However, further study is required of how to deal with this unstable region in tetrachromatic matching. Stiles (1955) added green to the violet test colour for greater stability in trichromatic matching and Fry (1957) also drew attention to the problem.

In all the above tests of the uniqueness of the tetrachromatic match, the 4 compared curves were obtained on different days. No results were rejected and then repeated, and no attempt was made of a day-to-day normalisation of results. The conventional method of representing trichromatic matches, involving normalisation which tends to mask session-to-session variation of the actual match quantities, does not yet have a tetrachromatic counterpart.

It appears from the results that a unique tetrachromatic match does exist for any given test colour and it can be reached from a favourable initial trichromatic match in about 4 iterations. The whole procedure of an iteration takes, on average 4 minutes—twice as long as the trichromatic match. Thus, a

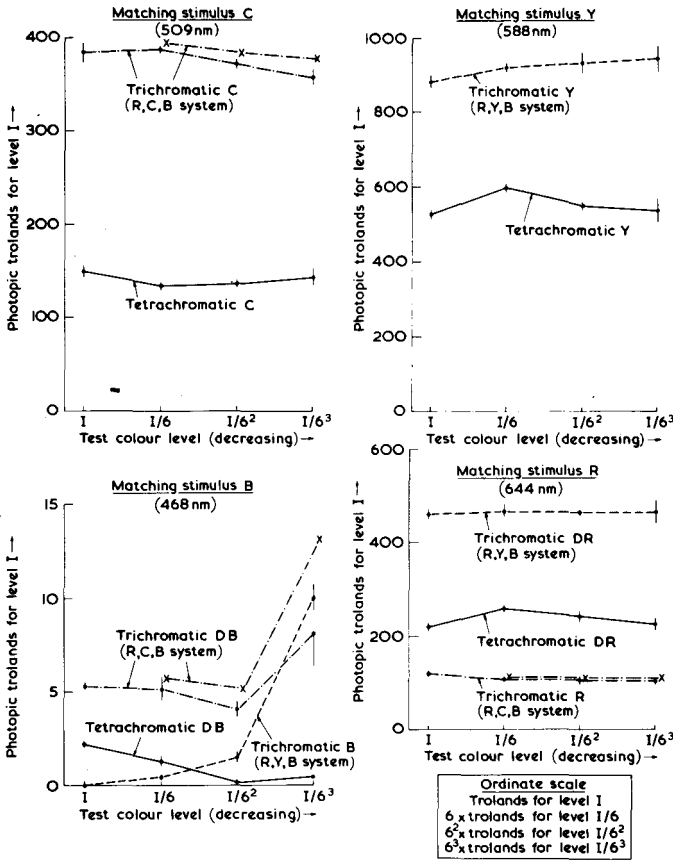


Fig. 4. Trichromatic and tetrachromatic colour matches at four levels. Test stimulus 546 nm, 520 photopic trolands at experimental level I. Lines show $\pm 2 \times$ standard error of mean. Crosses show a repeat at a slightly lower level for trichromatic matching (R, C, B system).

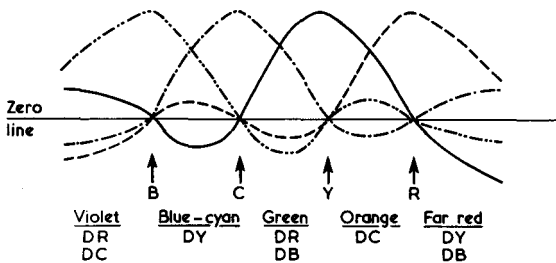


Fig. 5. Schematic tetrachromatic colour matching functions showing predicted desaturating stimuli in the five spectral regions.

tetrachromatic match can be reached, without prior knowledge, in about 15-20 minutes.

Fig. 3 shows for test stimulus 530 nm the effect of decreasing the luminance level on the tetrachromatic colour match and the trichromatic match (R, Y, B system). There are 4 graphs, one for each of R, Y, C, B , but C is absent in the trichromatic case. No match change with luminance level is indicated by a line parallel to the abscissa and its departure from this position is an indication of changes with level. Each point is the mean of 5 readings. Any change must be judged in conjunction with match uncertainty, and so lines show the limits of $\pm 2 \times$ standard error of the mean, i.e. another mean has a 95% probability of lying within these limits provided no changes have occurred. Hence point variations of this magnitude are not significant. Fig. 4 shows, for test stimulus 546 nm, tetrachromatic matches at different luminance levels, together with matches for 2 trichromatic systems — R, Y, B , and R, C, B . PGM was the observer for both test stimuli. In each case the R, Y and C matching stimuli probably show no significant changes. Large *increases* of the proportion of B with diminution of level occur in each trichromatic case while there is either no change or a *decrease* of the proportion of B in the tetrachromatic case (discussed later). Precision, as indicated by the shortness of lines in Figs 3 and 4, is rather worse in tetrachromatic than in trichromatic matches.

The changes in the blue matching stimulus for test stimulus 546 nm were confirmed rather differently for another observer *ERF* who made a match at the level next to the highest level and then viewed it at the lowest of the 4 levels. For both the R, Y, B and the R, C, B systems the trichromatic match was not acceptable. However, it could be made completely acceptable in each case by increasing the appropriate blue matching stimulus only and not touching the other controls. On the other hand the tetrachromatic match was quite acceptable at the lower level. Moreover, it was viewed at all levels between the experimental level and the region of the threshold of vision in $\times 2$ steps (20 levels in all) and no mismatch could be detected. *ERF* was also unable to detect a change at the lower level for other test stimuli, one being taken from each of the 5 spectral regions bounded by R, Y, C and B .

DISCUSSION

To predict the observed change from trichromatic to tetrachromatic match, let us consider "No initial R " for an orange test colour (Fig. 2*b*). At stage 0 we should expect the complementary colour C to be negative, i.e. DC with Y and B positive; R has been made zero. The orange- C field probably has a greater scotopic content than $Y-B$, so at stage $\frac{1}{2}$, DC will be reduced. The colour of the field is affected and cyan must be removed from the mixture field at stage 1. But C is equivalent to positive Y and B with negative R ; so Y and B must be

decreased while R is increased in the mixture field.

It is also possible to predict the desaturating stimuli for each of the five spectral regions. Fig. 5 shows schematic *cmfs* which are meaningful only in the way they cross the zero line. At wavelength R , the other matching stimuli are zero, and the same applies for Y , C and B , and Fig. 5 also assumes that they are never zero elsewhere. It can be seen that the number of *DSs* alternates between 1 and 2 in adjacent spectral regions; each matching stimulus occurs in 2 regions. Those of us who have thought of a common violet-magenta-deep red region, all requiring the adjacent R and B as mixture stimuli and G for a desaturating stimulus, must revise our ideas. This is purely a consequence of the trichromatic system; in the tetrachromatic system, the desaturating stimuli are completely reversed in the deep red and violet spectral regions. The desaturated colour investigated required amounts of all four matching stimuli to be positive.

Another way in which the tetrachromatic match could have been approached is by starting with C at the "low level", followed by R , Y , B at the experimental level. Adjustment of the "low level" is not quick, as changing the sector angle entails switching the motor off and on, and unclamping and clamping the disc. Starting at the "low level" may result in big changes of angle being necessary, each change requiring a few trials. For test colours in the yellow, orange and red regions, there is little or no scotopic activity to balance until the other matching stimuli, especially B , are introduced. Thus a big aperture would be required at first, with a smaller one after the first experimental level match. Fig. 2 shows that starting at the experimental level with R , Y , B , the matching stimuli subsequently used at this level, shows no advantage over the other 3 starting conditions.

The tetrachromatic match, by its nature, is less precise than the trichromatic, being dependent on a scotopic luminance match; careful choice of "low level" for each test colour may well yield greater precision. However, it is not only the cyan matching stimulus that is imprecise since changes in C are compensated for by changes in R , Y , B as described above. Taking Y at a shorter wavelength may help in improving precision in the experimental level match; but if Y is too close to C , the match might converge too slowly or even diverge. Just how wrong it is possible to make the approach to the final match has been shown by reversing C and B , ie by using B at the "low level". Divergence was so rapid that all energies had increased by a factor of 40 or more by stage 2.

With decreasing luminance a big increase occurs in the amount of blue matching stimulus for trichromatic matches on some test stimuli, but no change or else a decrease occurs for the tetrachromatic case. For instrumental reasons a matching stimulus can give too low a reading in the region of its colour threshold (Trezona, 1954), a problem usually involving the B component of a colour match. With a density wedge, a just discriminable step is usually about

the same size above and below the match and so the bracketing technique yields the "correct" reading. But near the colour threshold the step below the "correct" match is much larger than that above and so the bracketing technique on a density wedge gives too low a reading. Consequently, a *decrease* in a small quantity at low levels, as seen for *DB* in the tetrachromatic match, can have an instrumental cause rather than one of match failure. But an *increase* cannot be explained by such an experimental artifact. For 546 nm and the trichromatic *R, Y, B* system, the blue in the mixture field increases with decreasing level; calculations based on the V'_λ curve indicate that the test field at level I has a scotopic luminance $3\frac{1}{2}$ times that of the mixture field. For the trichromatic *R, C, B* system, it is *DB* that increases with decreasing luminance level, while the mixture field at level I has a scotopic luminance $2\frac{1}{2}$ times that of the test field. Both are in keeping with the suggestion that excess rod activity is balanced by the blue matching stimulus (Trezona, 1970).

Finally, we can consider the question of where a metric, based on the tetrachromatic match with its additive properties, could lead us. That is, how can a 4-dimensional system be successfully applied to a situation where colours are completely specified in 3-dimensions? The ultimate aim of the investigation is to enable a correct chromaticity to be calculated at any luminance level, given an absolute spectral energy distribution under standard conditions. It needs to be *absolute* since chromaticity is luminance-dependent under large field conditions. The tetrachromatic match enables the first part of this problem to be solved since, being additive, it provides a system where summation can be performed validly. The second part of the problem, which the author hopes to tackle, consists of emerging, once summation is complete, from the tetrachromatic into a trichromatic system in a way which depends on the luminance level.

ACKNOWLEDGMENTS

I am greatly indebted to Dr. F. J. J. Clarke for many helpful discussions. Thanks are due to Mr. Peter Moss and Mr. Ed. Fulton for acting as observers. This work forms part of the research programme of the National Physical Laboratory.

REFERENCES

- Bongard, M. M., et al. (1957), The four-dimensional colour space of the extra-foveal retinal area of the human eye. Proc. Symp. 'Visual problems of colour', Vol. I. Teddington, 325-330
- Clarke, F. J. J. (1960), Extrafoveal colour metrics. Opt. Acta 7, 355-384
- Clarke, F. J. J. (1963), Further studies of extrafoveal colour metrics. Opt. Acta 10, 257-284
- Clarke, F. J. J., and Trezona, P. W., to be published.
- Crawford, B. H. (1949), The scotopic visibility function. Proc. Phys. Soc. (Lond.) B 62, 321-334
- Crawford, B. H. (1965), Colour matching and adaptation. Vision Res. 5, 71-78

- Crawford, B. H. (1962), in: H. Davson (ed.), *The eye*. Vol. II New York, etc., Academic press, p. 68
- Fry, G. A. (1957), Chromatic adaptation with special reference to the blue-green region of the colour-mixture diagram. *Proc. Symp. 'Visual problems of colour'*, vol. II Teddington, 665-680, 733
- Speranskaya, N. I. (1959), Determination of spectrum colour co-ordinates for twenty-seven normal observers. *Optics. Spectrosc.* 7, 424-428
- Stiles, W. S. (1955), 18th Thomas Young Oration: The basic data of colour-matching. *Phys. Soc. Year Book*, 44-65
- Stiles, W. S., and Burch, J. M. (1959), NPL colour-matching investigation; final report (1958). *Opt. Acta.* 6, 1-26
- Trezona, P. W. (1954), Additivity of colour equations II. *Proc. Phys. Soc. (London) B* 67, 513-522
- Trezona, P. W. (1970), Rod participation in the 'blue' mechanism and its effect on colour matching. *Vision Res.* 10, 317-332

DISCUSSION

Wyszecki: I suppose the particular choice of primaries used will have some effect on the speed of achieving a match at both luminance levels?

Trezona: Yes, this is true. In one test I changed over the blue and cyan matching stimuli using the blue at low level and the cyan at high level, expecting I might get slower convergence. Instead I found divergence; so the choice of these is very important, especially that of the cyan used to balance rod response. I suspect that it cannot be taken very far from the peak of the V'_λ curve.

Hunt: Why was the blue primary not chosen to be at a shorter wavelength than 468 nm, so as to reduce its excitation of rods?

Trezona: Yes, this could in fact be done but there is no convergence difficulty in any reasonable spectral location for the blue primary. As long as the big changes occur initially (Fig. 2b, e.g.), this does not matter too much, but I agree that you could get faster convergence probably by taking blue at a shorter wavelength.

Wyszecki: There may also be sets of primaries where no "convergence" may be attainable e.g. if the *C* and *B* primaries are very close.

Trezona: Indeed there will be combinations where you get very slow convergence and other combinations where you would get divergence. We never get divergence for that particular set of four matching stimuli. I think there is a considerable range over which they could be chosen, but possibly the choice of *C* would be much more critical than the choice of the others. Yes I agree you could probably chose a set of four which would not converge or would converge more slowly than would be practically useful.

Stiles: Now what next?

Trezona: Well the next thing I want to do is to put the whole thing on a firmer basis and as I said, to determine exactly how to fix the low level. This has been left very largely to the observer up to now. He objects if the level is too high or too low, but I want this defined rather better, and I also want to see the effect of changing matching stimuli. Having decided on these things I can

3 DISCUSSION

then measure the tetrachromatic color matching functions. Lastly I could make a trichromatic match between the four matching stimuli and then at high levels the tetrachromatic color matching functions should reduce to trichromatic ones and a comparison could be made.

McLaren: Is this work relevant to the matching of surface colours? Will luminance metamerism occur under these circumstances?

Trezona: If we could have four photocells with, as response, linear transformations of the tetrachromatic colour matching functions and if two surfaces gave the same four readings not only would they match in colour but they would not show luminance metamerism.

McLaren: You don't answer my second question. At the levels used for surface colour matching, is the effect important?

Trezona: Well, I think from speaking to other people that luminance metamerism is a problem. Maybe somebody else could answer this problem.

Wyszecki: In reply to Mr. McLaren's question: In large-field viewing situations the rod mechanism may not completely be eliminated from functioning even at very high luminance levels. When strong metamerism is present, particularly for red colors, the intrusion of the rods may be significant and make match-predictions by a trichromatic system (such as the CIE large field system) incorrectly.

Malkin: We had an example of components for a bathroom suite which were an adequate match at high levels of illumination, but not at lower levels e.g. ambient light in a bathroom on a dull day. No significant metamerism was noted on changing from daylight to tungsten light at high levels of illumination.

Clarke: The consumer has not got a powerful Xenon arc to do his viewing by! If the colour finishing industries regard the consumer's judgment as important then luminance metamerism must be a relevant and practical problem. The availability of daylight via window apertures is very limited in winter months in many parts of the world, and rather high levels are needed to eliminate rod activity.

Jaekel: I support Dr. Clarke's remarks. During the winter months we often had to send away assessors who had chosen to use daylight for "pass-fail" decisions, because the light-intensity was insufficiently high for critical judgments.

McConnell: The question of luminous metamerism was mentioned for paper colours. Unfortunately the papers were of high reflectance say 90-100% and were viewed under conditions of good illumination. A possible explanation was given at Stockholm under a revised neutral axis for assessing near white papers.

Reilly: In response to Mr. McLaren's question: Dr. Trezona's data seem to show that lightness metamerism could become a problem with surface colors whenever the illumination and reflectance combine to produce a retinal il-

luminance about 1/36th the photopic reference level used in her experiments.

Ganz: At which levels do things become relevant? Would the influence of rod vision be effective at the illumination level in this lecture room—or are we not perhaps looking for metamerism in the dark?

Trezona: Well, it starts in the photopic region, but the mesopic region is not very far below. The limits of the mesopic region are ill defined.

McLaren: If two blacks are being matched in a brightly illuminated color matching cabinet, is rod intrusion likely?

Trezona: Yes, I should think so. But this would depend on the field size and the reflectance of the samples as compared to the background.

Lozano: The problem of non-additivity in large fields has been proved. In the light of it I think that Dr. Trezona's work can be useful to find an answer on this matter. If rods have something to do with large-field color vision, perhaps in this way it is possible to find how.

Simon: In industrial conditions for matching, one seeks the best lighting one can get, and uses high levels. We are not attempting really difficult matches when the level is extremely low.

Schanda: In case you are performing your investigations at different levels of illumination, and are using different adaptation times, it could be perhaps possible to find answers on the questions raised concerning rod and cone vision.

Trezona: I have not changed the adaptation time but I don't think it would affect the tetrachromatic match because this is a *physiological identity* in the two parts of the field. Between high and low illumination we did not allow much time; I think it might affect the absolute level of the low level but not the actual amount of matching stimulus used, because of the identity in receptors (rods only) between the two parts of the field.

4 WAVELENGTH DISCRIMINATION MEASURED BY CHROMATIC FLICKER STIMULI

NAOTAKE KAMBE

*Toshiba Research and Development Center,
Tokyo Shibaura Electric Co., Ltd., Kawasaki, Japan*

INTRODUCTION

Discrimination of wavelength differences has usually been studied under various conditions of field size, retinal illuminance, position of the retina and color vision of the subject. Wavelength discrimination curves measured by Willmer and Wright (1945), Bedford and Wyszecki (1958) and McCree (1960) were obtained under conditions of variable field size and viewing technique. The data presented by Willmer and Wright show that decreasing the field size results in a tritanopic confusion in the blue-green part of the spectrum. They employed the technique of strict foveal fixation on the visual field. Bedford and Wyszecki used the technique in which the subject was allowed to continuously scan the field, and found little or no tritanopic confusion.

The curves measured by McCree, however, show a tendency to a tritanopic confusion with decreasing the field size, although the subject was allowed to scan the field in order to avoid strict fixation. Bedford and Wyszecki have suggested that these differences in wavelength discrimination curves might be attributed not only to the different viewing techniques used but also to the difference in type of visual field. The curves presented by McCree were obtained using the bipartite field with horizontal dividing line, whereas those presented by Bedford and Wyszecki were obtained using two circular points separated vertically.

In the experiments reported here, wavelength discrimination curves were obtained using another different type of visual field, that is, the curves were determined using the chromatic flickering field of alternating one spectral light and another spectral light at equal luminance.

There are several experiments concerning the dependence of critical chromatic flicker frequency on the magnitude of the chromatic differences of the flick-

ering components, for example De Lange (1958) and Van der Horst (1969). However, these results are limited to a few pairs of spectral lights. Only Walraven, Leebeek and Bouman (1958) measured the fusion frequency of alternating white and mixtures of white and spectral light as a function of saturation and wavelength. To the author's knowledge, no information was found in the literature on the wavelength discrimination throughout the spectrum as a function of chromatic flicker frequency. The experiments were designed to examine the role of the variables of chromatic flicker frequency and field size in determining the least noticeable differences of wavelengths.

APPARATUS

The optical system that produces the sinusoidally alternating colored stimuli is shown in Fig. 1. The method used is essentially the same as those used by De Lange (1958), slightly modified to meet the Maxwellian view system.

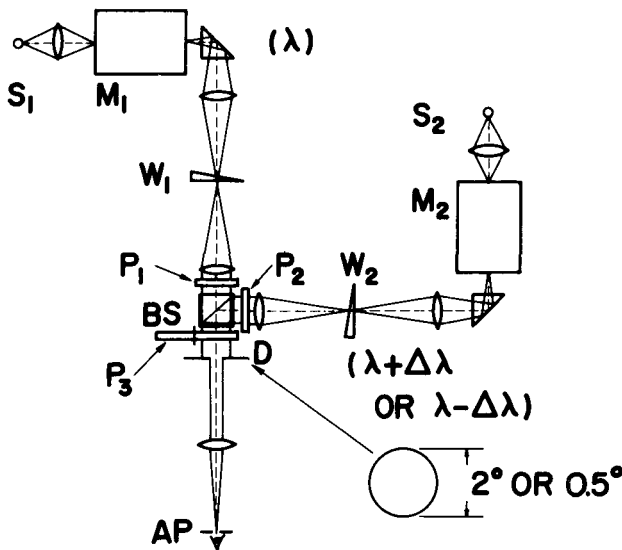


Fig. 1. Apparatus and appearance of the visual field.

The beam from a double prism monochromator M_1 is used as the reference light and the beam from a single grating monochromator M_2 as the comparison light. These two beams are sinusoidally modulated by means of the pair of fixed polarizers P_1 and P_2 and the rotatable polarizing disk P_3 . Placing the polarization axes of the polarizers P_1 and P_2 mutually perpendicularly, the two sinusoidally modulated beams have a phase difference of 180° , as shown in (I) and (II) in Fig. 2. The lower part (III) in Fig. 2 represents the temporal configuration of two alternating color stimuli when the luminance level of both lights is balanced with control of the neutral density wedge W_2 .

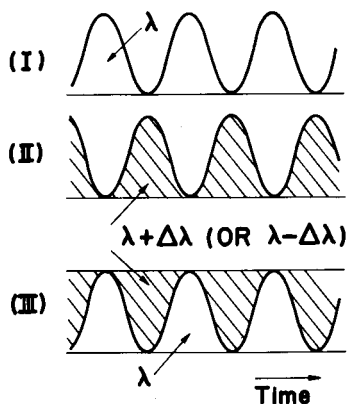


Fig. 2. Temporal configuration of the stimuli.

The color flicker frequency can be continuously controlled by varying the speed of the rotatable polarizing disk P_3 with the change of the input voltage to an alternating current servo motor. The field size can be varied with the change of the diaphragm D .

The monochromators were calibrated with the lines in mercury, cadmium and sodium spectrum. The half-peak bandwidth was fixed at 10 nm throughout the experiments. A blue filter (Toshiba VV-40) was used to block the stray light from the single grating monochromator at the measurement between 440 nm and 400 nm.

PROCEDURE

The experimenter first adjusts the reference and comparison lights to the same wavelength and sets the luminance of reference light. The subject varies the luminance of comparison light until the luminance flicker is minimal, possibly zero. The subject then gradually changes the wavelength of the comparison light to one side until a difference of wavelengths is reached at which he can not make the flicker disappear by adjusting the luminance of the comparison light. The detected flicker is then due to the change in wavelength. Fixation point was not presented to the subject, and he was allowed to freely scan the visual field. Similar settings were repeated twice on both sides of the wavelength.

This procedure was repeated four or six times for each wavelength at 20 nm intervals from 400 nm to 620 nm. The alternating frequency was varied in three steps of 1, 3 and 5 Hz, and the field size in two steps of 2° and 0.5° in diameter. The retinal illuminance of the circular field was approximately 20 trolands for all conditions.

The subjects were two males of normal color vision, subject *NK* in his early thirties and *JA* in his early twenties.

RESULTS AND DISCUSSION

Wavelength discrimination curves for two subjects are shown in Figs. 3 and 4. The magnitude of least noticeable wavelength differences in the direction of longer wavelength, $\Delta\lambda$, and the magnitude in the shorter wavelength, $|\Delta\lambda|$, are plotted against the wavelength $\lambda + \Delta\lambda/2$ and $\lambda - |\Delta\lambda|/2$, respectively. The continuous lines through the plotted points are drawn to give an indication of the general trends of the wavelength discrimination curves.

These curves show that increasing the alternating frequency results in poorer discrimination for the whole range of the spectrum, as is to be expected. Much deterioration of discrimination also occurs in going from the 2° field to the 0.5° field. There seems to be no significant shift of the locations of relative maxima and minima on the curves with the changes in alternating frequency and field size. The positions of two relative maxima and three relative minima are generally the same as those previously obtained with the bipartite field.

However, the shape of the curves does not retain the same shape when plotted on a logarithmic scale. It is found that the shapes of the curves are more sensitive to the field size than to the alternating frequency. The curve of the 0.5° field for both subjects show much deterioration of discrimination, particularly in the blue-green part of the spectrum. This selective lowering of blue-green discrimination is the same direction as a tritanopic wavelength confusion. In the case of subject *NK*'s data of the 0.5° field, increase of the alternating frequency from 1 Hz to 5 Hz also produces much deterioration in the blue-green region. This type of change appears fairly in subject *NK*'s data of the 2° field and appears slightly in subject *JA*'s data of the 2° field, although the change does not appear in the curves of the 0.5° field for subject *JA*. The results may be interpreted as an indication that increasing of the alternating frequency has an effect similar to fixating strictly on the visual field with respect to wavelength discrimination.

Walraven and Bouman (1966) have suggested that normal color vision can be specified as a combination of tritanopic and deuteranopic vision. They have assumed that reducing of the contribution to color vision of the yellow-blue system leads to a tendency toward tritanopic vision. On this hypothesis, the results may suggest that the variable of the field size is more effective to the contribution to color vision of the yellow-blue system than that of the alternating frequency.

Further experiments are needed on the chromatic discrimination in non spectral purple region, because it can be expected that the chromatic discrimination may become especially poor in this region with reducing of the contribution to color vision of yellow-blue system. Other checks are also needed on the effect of brightness flicker which is due to the phase shift between the responses in the visual system shown by de Lange (1958).

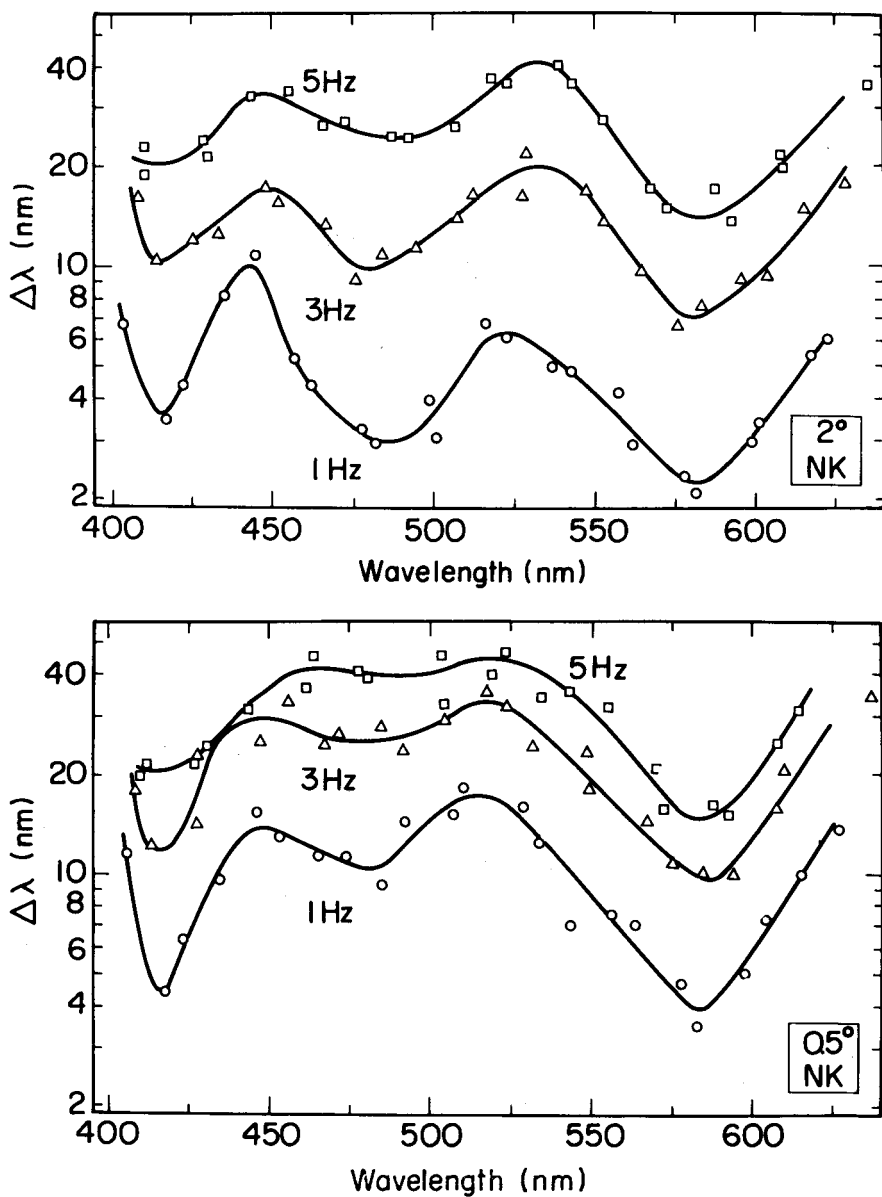


Fig. 3. Wavelength discrimination curves for subject NK for two field sizes and various alternating frequencies at 20 trol.

ACKNOWLEDGMENTS

The author wishes to thank Dr. L. Mori for his continuous encouragement and advice, Messrs. J. Akiyama and Y. Tanaka for their help throughout the experiment.

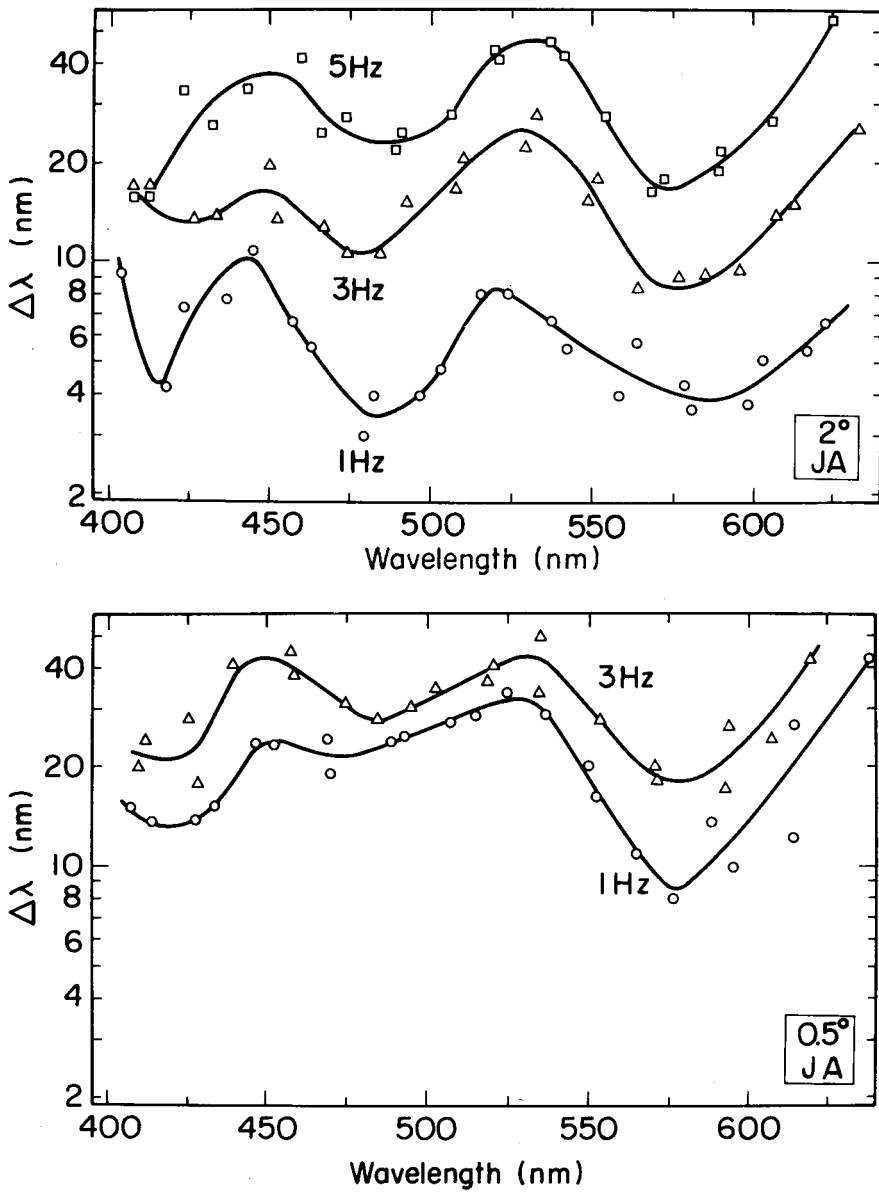


Fig. 4. Wavelength discrimination curves for subject JA for two field sizes and various alternating frequencies at 20 trol.

REFERENCES

- Bedford, R. E., and Wyszecki, G. W. (1958), Wavelength discrimination for point sources. *J. Opt. Soc. Amer.* 48, 129-135.
- McCree, K. J. (1960), Small-field tritanopia and the effects of voluntary fixation. *Opt. Acta* 7, 317-323.

4 DISCUSSION

- Horst, G. J. C. van der (1969), Chromatic flicker. *J. Opt. Soc. Amer.* 59, 1213-1217.
- Lange Dzn, H. de (1959), Research into the dynamic nature of the human fovea-cortex systems with intermittent and modulated light. II. Phase shift in brightness and delay in color perception. *J. Opt. Soc. Amer.* 48, 784-789.
- Walraven, P. L., and Bouman, M. A. (1966), Fluctuation theory of colour discrimination of normal trichromats. *Vision Res.* 6, 567-586.
- Walraven, P. L., Leebeck, H. J., and Bouman, M. A. (1958), Some measurements about the fusion frequency of colors. *Opt. Acta* 5, 50-54.
- Willmer, E. N., and Wright, W. D. (1945), Colour sensitivity of the fovea centralis. *Nature* 156, 119-121.

DISCUSSION

Clarke (Chairman): Perhaps, since Dr. Kambe finished before time, may I be permitted to give a slight lead into the discussion: the conclusion on this subject of the importance of field size is the same as can be inferred from the work in Brindley and Stiles using the two colour approach and some of the work I did a few years ago at the University of Rochester.

Wagner: Do you have slides, that show the comparison between your experimental results and the theoretically predicted curves?

Kambe: Judd and Yonemura (1970) proposed an opponent model on color vision. Recently, Yonemura (1970) proposed a revision of this model. In this model, the theoretically predicted wavelength discrimination curves of normal vision are constructed from the results of combining the curve of tritanopic vision with that of protanopic (Judd and Yonemura) or deutanopic (Yonemura) vision. The theoretical curves change gradually toward the tritanopic curve with reducing of the contribution of the yellow-blue component. I have made some comparisons of our experimental curves with this type of theoretical curves. The degree of agreement is fairly good, including the minimum in the violet part of the spectrum.

Valberg: Why did you use a logarithmic scale for plotting $\Delta\lambda$? Was it just convenient or had it theoretical significance?

Kambe: The logarithmic plot is convenient to show proportional variations of wavelength discrimination curves. If nothing happens but a shift on a logscale, we know that only one factor has to be changed.

Kowaliski: Dresler (1953) demonstrated that flicker measurement techniques yield results identical with those obtained with instruments (barrier-layer or photoemissive cell responses) whereas static matches correspond to visual (eye sensitivity) responses. In spite of the low frequencies used (1, 3 and 5 Hz), some doubt may arise about the relationship of these results with eye responses.

Kambe: In the case of heterochromatic flicker photometry, a subject changes the luminance of comparison light until the brightness flicker is minimal above the chromatic fusion frequency. But, in the method of observation we adopted, a subject changes both the wavelength and luminance of comparison light until

the wavelength difference produces a just perceptible color flicker when the brightness flicker is minimal. I think, we can get some information about the temporal characteristics of mechanisms on wavelength discrimination using this type of the experimental procedures.

REFERENCES

- Dresler, A. (1953), The non-additivity of heterochromatic brightnesses. *Trans. Illumin. Engng Soc.* 18, 141-165
- Judd, D. B., and Yonemura, G. (1970), CIE 1960 UCS diagram and the Müller theory of color vision. *J. res. Nat. Bur. Standards* 74A, 23-30
- Yonemura, G. (1970), Opponent-color-theory treatment of the CIE 1960 (u,v) diagram: Chromaticness difference and constant hue loci. *J. Opt. Soc. Amer.* 60, 1407-1409

5 SIMILARITY BETWEEN JND-CURVES FOR COMPLEMENTARY OPTIMAL COLOURS^{*})

A. VALBERG AND T. HOLTSMARK^{**}

*Laboratory of Colour Metrics,
Physics Institute, Basel, Switzerland*

INTRODUCTION

Recently we have described a simple lecture room experiment, in which pairs of mutually complementary colour stimuli distributions are projected on a screen (Holtsmark and Valberg, 1971). The stimuli distributions derive from pairs of mutually complementary apertures (Fig. 1), which are projected through a prism.

Complementary colours are opposing each other in the two "spectra", and it appears that rapid change of colour occurs at corresponding positions in the spectra. This circumstance raises the question about a possible correspondence between discrimination curves for pairs of complementary colour stimuli distributions.

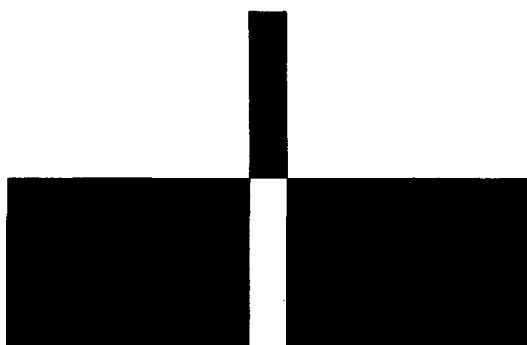


Fig. 1. A "positive" and a "negative" aperture giving rise to complementary colour stimuli distributions.

^{*} This work has been supported by Grant 2.173.69 to K. Miescher from Schweiz. Nationalfonds zur Förderung der wissenschaftlichen Forschung and by a grant to T. Holtsmark from Norges Almenvitenskapelige Forskningsraad.

^{**} On leave from University of Oslo, Physics Institute.

METHOD AND APPARATUS

This report is concerned with an extension of earlier measurements of just noticeable colour differences (*JDN*) for pairs of optimal colours, complementary with regard to a white surround (Holtsmark and Valberg, 1969).

The colour stimuli were established by means of a Spectral Colour Integrator developed at this laboratory (Gasser et al, 1959; Weisenhorn, 1965) (Fig. 2).

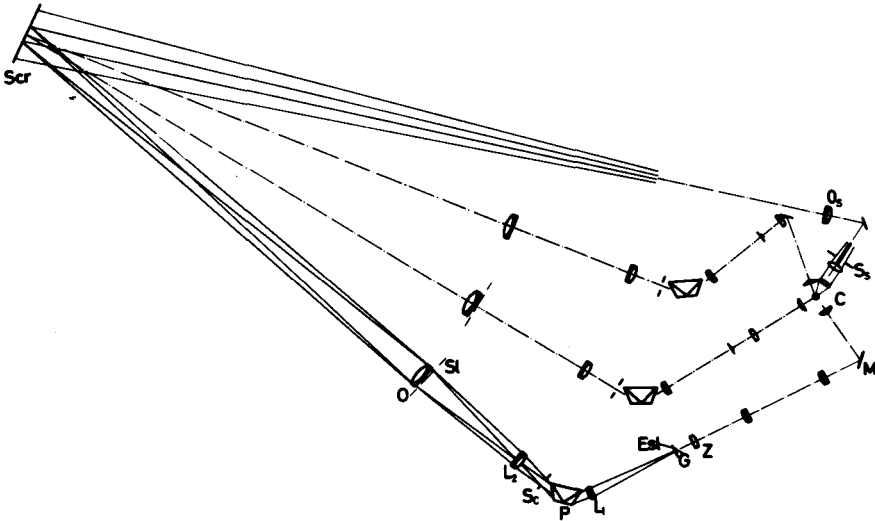


Fig. 2. The principle of the Spectral Colour Integrator.

The main principle of the instrument is a combination of two image formations. The entrance slit (*Es*), which is illuminated by the collimated light beam from a 1600 W High Pressure Xenon Arc source, is imaged as a 100 mm broad prism spectrum at the plane (*Sl*). A field stop (*Sf*) before the prism (*P*) is imaged by the objective lens (*O*) on a $BaSO_4$ -screen (*Scr*) at a distance of 2.5 meters in front of the instrument. The colour of this image can be deliberately determined by masking off parts of the spectrum by means of moveable apertures (*Sl*). In our experiments, we have used only rectangular "positive" and "negative" apertures (± 10 mm, ± 20 mm, ± 60 mm) which give rise to complementary optimal colour stimuli. (For a discussion of optimal colours, see Bouma, 1951). The chromaticity loci for various slit widths are shown in Fig. 3.

Two prism channels formed the bipartite comparison field, consisting of two contiguous crescents totalling 2.3° , while a third, nondispersing channel projected the white 13° surround of 180-200 *asb* luminance. Identical positions of identical apertures (*Sl*) in the two spectra give rise to matching colours in the bipartite field. By moving one aperture toward shorter or longer wavelengths, a position will be reached where the two colours become different.

To enable measurements on a complete colour cyclis, a double slit (*Sl*) was applied. When the edge of one slit passes out of the spectrum at 700 nm, the opposition edge of the other slit enters at 400 nm.

The apertures were mounted on motor driven slides. One aperture was operated by the experimenter who thus determined the colour of the left field which served as reference. The other slit was operated by the observer. He could also adjust the motor velocity.

For each setting of the left reference field, the observer adjusted the colour of the right

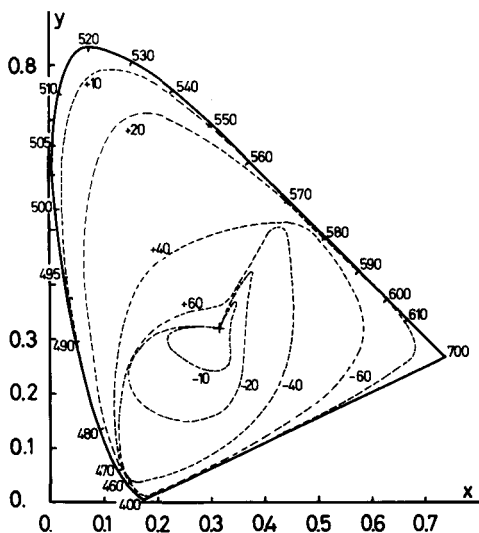


Fig. 3. Chromaticity loci for complementary optimal colours, slitwidth in mm serving as parameter.

field to establish six matches and six differences, three to each side of the match. Between each of the twelve adjustments, the central field was blackened out for at least 30 seconds or until the observer did not perceive any after-image in the central field. The observer had to establish a just readily noticeable difference. He was allowed to look freely over the visual field. No artificial pupil was applied. If necessary for establishing the first match of a standard setting, the observer equalized the lightness. Apart from this, no equalization of the lightness was introduced. The observer did not distinguish between different colour attributes, he just responded to a difference. If he felt uncertain about the difference or found it necessary to move the slit back and forth, the adjustment was interrupted and a repeated 30 seconds adaptation time was introduced. The time used in setting up a match or a difference proved to be important. It should not exceed three seconds.

After the sequence of twelve adjustments, the experimenter moved the reference aperture a distance corresponding to about two *JND*. One experimental session generally consisted of nine standard settings. Between successive sessions, a resting period of 30 minutes was introduced.

REPRESENTATION OF RESULTS

A representation of the displacement of the apertures in mm delivers a critical test of the symmetry of discrimination for complementary stimuli distributions. In Fig. 4, the displacement Δs represents the mean of the average differences Δs^+ and Δs^- to each side of a match measured by the position of the right edge:

$$\Delta s = \frac{1}{2} (|\Delta s|^+ + |\Delta s|^-)$$

To represent the thresholds for positive and negative slits on the same wavelength scale, for positive slits, the difference $\Delta \lambda_a$ of dominant wavelengths has been plotted against λ_a

$$\Delta \lambda_a = \frac{1}{2} (|\Delta \lambda_a|^+ + |\Delta \lambda_a|^-).$$

For negative slits, the compensative wavelength λ_c and the difference $\Delta\lambda_c$ have been computed accordingly. (For computer programming see Richter, 1965).

The solid lines below the + 10 mm, + 20 mm and + 60 mm curves mark the wavelength range covered by the slitwidth for two or three selected dominant wavelengths of the colour stimulus. Due to the nonlinear dispersion of the prism, the extension of this range changes from one part of the spectrum to another.

Because we are simulating object colour conditions, a representation in an object colour space of the Luther-Nyberg type may prove convenient (Luther, 1927). Furthermore, in this space complementary colours are symmetrically spaced about the center. The co-ordinates of this system are the partial chromatic moments M_1 , M_2 and the luminance factor Y . Provided the white colour stimulus has the *CIE* tristimulus values $X = Y = Z = 100$, the co-ordinates are linear transformations of the *CIE* co-ordinates according to an opponent colours scheme:

$$\begin{aligned}M_1 &= X - Y \\M_2 &= k(Y - Z), \text{ } k \text{ is a constant} \\Y &= Y \\M &= [M_1^2 + M_2^2]^{\frac{1}{2}}\end{aligned}$$

Strict similarity of the *JND* curves for positive and negative slits would imply that, in the linear colour space, the *JND* vectors have the same absolute magnitude for complementary colour stimuli. This magnitude is given by

$$\Delta E = [\Delta Y^2 + \Delta M_1^2 + \Delta M_2^2]^{\frac{1}{2}}$$

In Fig. 5, ΔE is represented for obs. *TH*, for the ± 20 mm slitwidth series.

COLOUR DIFFERENCE FORMULAE

We have tested some of the best known colour difference formulae with respect to equal distance and symmetry. In Fig. 5, the solid curve represents *JND* for the ± 20 mm slits, computed as ΔE according to the formulae. For the complementary stimuli, the broken line represents ΔE as predicted by the formulae provided that the symmetry rule holds strictly.

DISCUSSION

The assessment of general colour differences far above threshold, faces the observer with the task of correlating more than two qualitative attributes. Judd (1968, 1969) has recently drawn attention to the theoretical problems involved in such kinds of judgements and in the related problem of constructing an equidistant colour space. Near the threshold level, however, where the

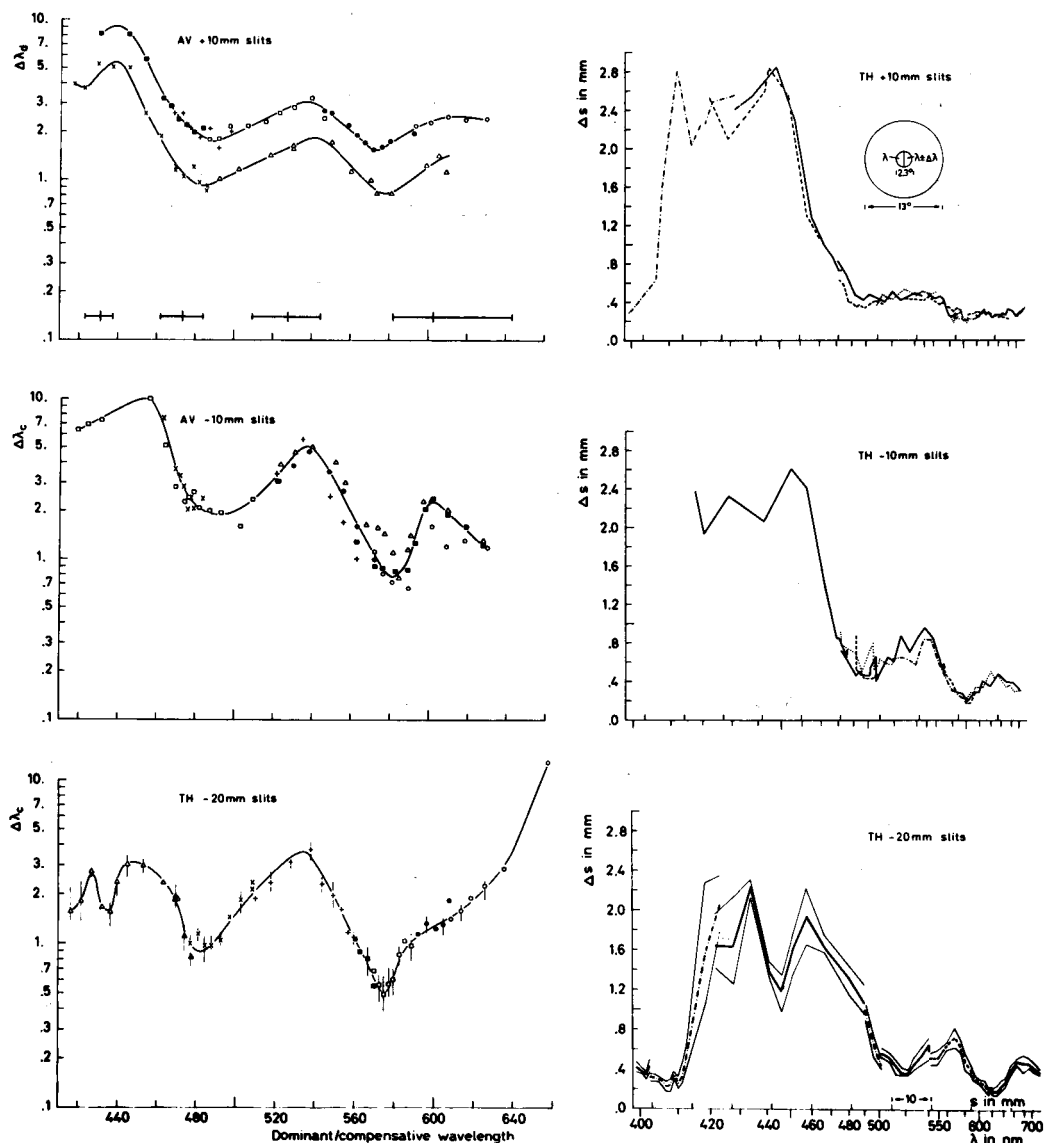
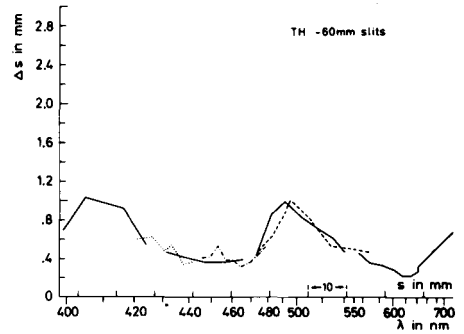
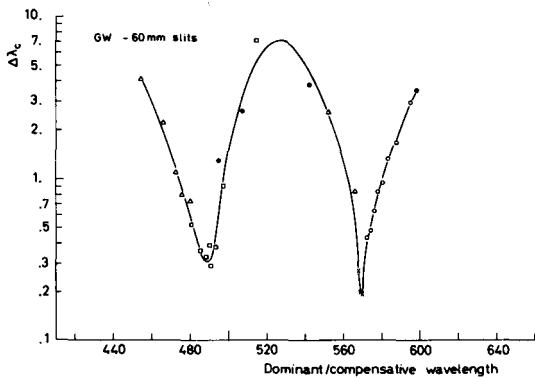
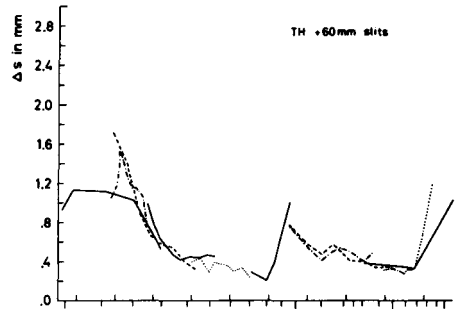
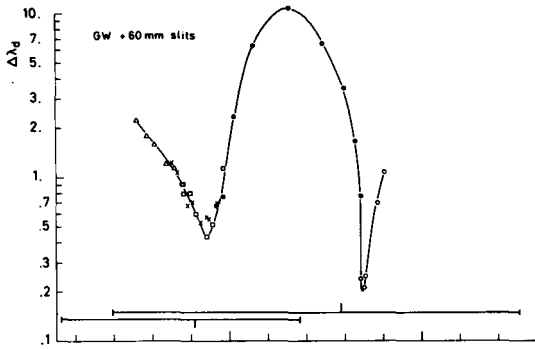
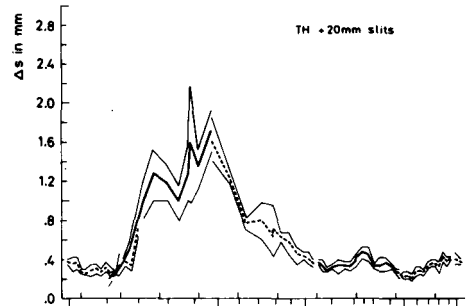
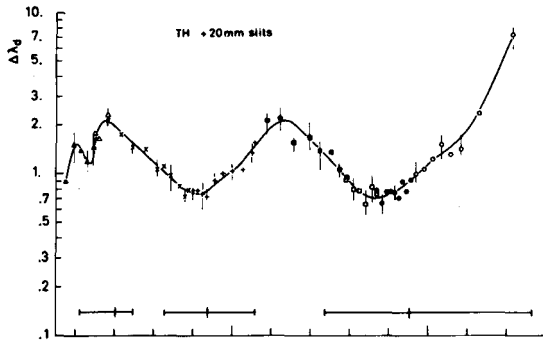


Fig. 4. Examples of symmetry between JND -curves for complementary optimal colour stimuli (three observers).

Left: λ_d and λ_c representation. Right: displacement of the slit in mm. For all ± 20 mm curves, obs. TH, the maximum deviations are indicated. Symmetry can be judged by comparing corresponding right-hand figures. Different symbols indicate different experimental sessions.



observer only decides on a difference, a unit discrimination step can be established without much difficulty.

With some training the observer was able to maintain a stable criterion of difference throughout a complete cyclis (usually lasting 4-5 days with 2-3 sessions a day) as will be seen from the maximal deviations shown in Fig. 4 ($TH \pm 20$). After a resting period of about three months, the criterion may have changed (Fig. 4, $AV, + 10$). We have also evidence that, above threshold, the observer is to a certain extent able to determine a degree of difference and

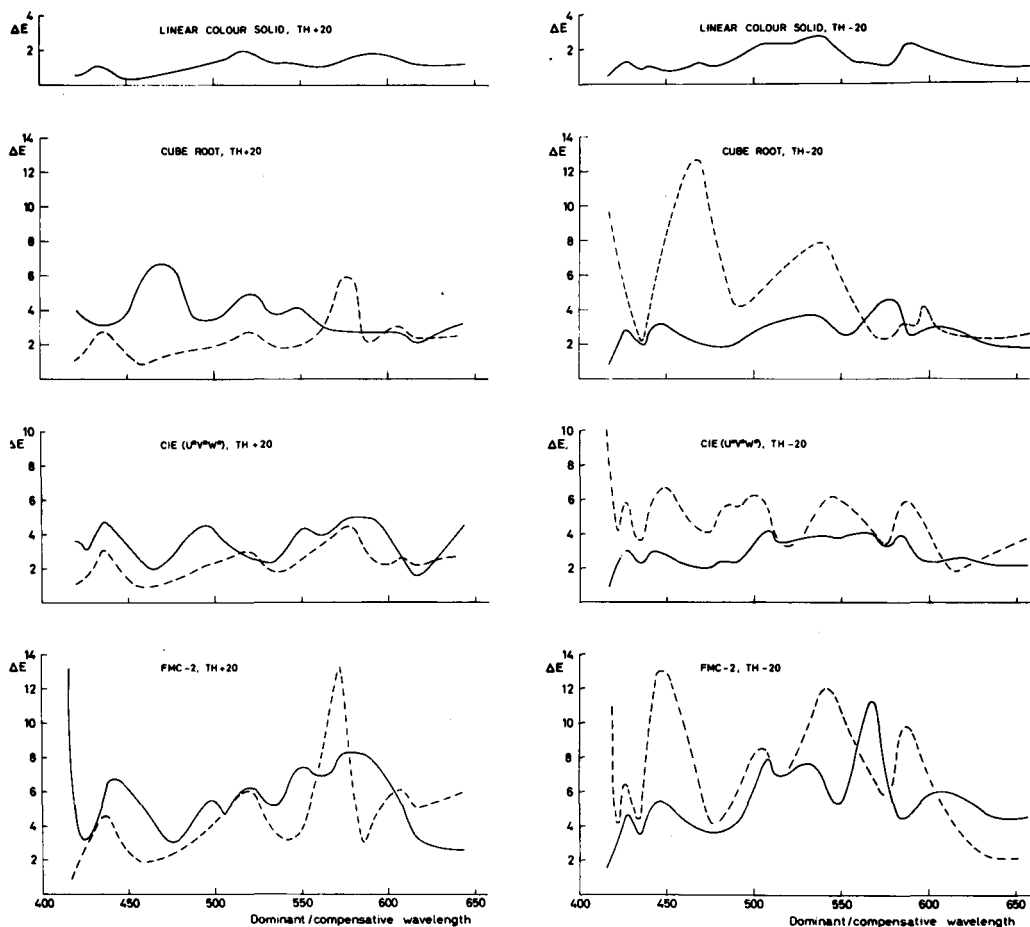


Fig. 5. JND represented as distances ΔE in linear colour space and as ΔE derived from three nonlinear colour difference formulae.

Broken curve indicates ΔE for the complementary stimuli as predicted on the assumption of symmetry.

keep to it for a whole cyclis. Curves representing various degrees of difference have the same shape and are shifted with a constant amount on a logarithmic scale.

In the linear representation of the position of the right edge of the apertures (Fig. 4), it generally appears that two neighbouring minima occur at a distance which equals the width of the aperture. The positions of the edges for these two minima are found in the 590-625 nm region. Thus, in this representation, a minimum occurs whenever an edge passes that region. A third minimum at the position 510-530 nm of the right edge does not seem to be duplicated according to the above mentioned principle.

For optimal colours, the stimuli differences at *JND* are obviously nearly the same for complementary colours. If one accepts our results as expressions for a symmetry relation, the cause for the deviation from this rule for narrow slits may be found in adaptation effects. *JND* curves for +10 mm slits determined in a dark surround showed a significant (40-50%) lower absolute level of $\Delta\lambda_d$ and Δs . Thus, with the white surround, the fact that differences are fairly equal for complementary colour stimuli might be at least partly explained by the maintenance of a rather constant adaptation level. Nevertheless, the central fields also contribute to the adaptation level. Therefore, it seems plausible that the deviation from symmetry should become most pronounced in the case of narrow positive and negative apertures (dark versus light colours).

It is seen from the Fig. 5 that when the *JND* for ± 20 mm slits are represented in a linear object colour space, the variation in the difference ΔE is smaller than for the more complex nonlinear formulae. Furthermore, the linear representation is symmetric. The same displacement of complementary slits represents exactly the same distance in the colour solid. Even the components ΔY , ΔM_1 and ΔM_2 (or ΔM) are equal. This last fact may have implications for the theory of interaction of component visual processes. At $\lambda_d = \lambda_c = 420$ nm $\Delta Y = 0.01$, and the fraction $\Delta Y/Y$ is about 0.2 for the +20 mm slit (relatively dark and saturated colours) and about 0.0001 for the -20 mm slit (light and unsaturated colours). Therefore, the same luminance difference gives rise to a greater lightness difference for the dark colours. Provided that symmetry is maintained, for the light colours, this smaller lightness difference must be compensated for by a greater hue and/or saturation difference.

In establishing a difference, simultaneous contrast seems to be an important factor. Although the observer adjusts only one half of the bipartite field, he perceives a simultaneous change in both fields. Without knowing, he will find it hard to tell which one of the two half-fields is actually adjusted. This fact indicates the existence of a close relationship between a threshold and a contrast process of the visual system (Valberg, 1971).

A crucial question in the theory of vision is concerned with the transformation of the retinal, trichromatic type of responses into opponent type of responses arising on some higher level of the visual pathway. On this level, any specific information about the spectral energy distribution of the stimulus has been lost. However, the organism is able to regroup the set of stimuli into mutual complementary pairs. Essentially, this is what happens through the opponent type of transformation. This might also explain why the difference sensitivity is of the same magnitude for complementary stimuli. A certain stimulus difference, being brought about by the displacement of a narrow (positive or negative) slit might equally well be reproduced by the appropriate displacement of a single edge. But in this latter case, the actual magnitude of the displacement of the edge at threshold is different from the threshold dis-

placement of the slit. Again, however, the threshold displacements for two complementary edges prove to be the same.

Thus, the difference sensitivity of the visual system seems to be closely attached to those response mechanisms which bring about the restructuring of the set of stimuli according to the law of complementarity.

REFERENCES

- Bouma, P. J. (1951), *Farbe und Farbwahrnehmung*. Philips, Eindhoven.
- Gasser, M., Bilger, H., Hofmann, K. D., and Miescher, K. (1959), *Spektraler Farbintegrator*. *Experientia* 15, 52.
- Holtsmark, T., Valberg, A. (1971), On complementary color transitions due to dispersion. *Amer. J. Phys.* 39, 201-204.
- Holtsmark, T., and Valberg, A. (1969), Colour discrimination and hue. *Nature* 224, 366-367.
- Judd, D. B. (1968, 1969), *Der ideale Farbraum. Palette* (Sandoz Basel) 29, 25-32; 30, 21-29; 31, 23-30.
- Luther, R. (1927), Aus dem Gebiet der Farbreizmetrik. *Z. techn. Physik* 8, 540.
- Richter, K. (1965), Programmierung farbmeterischer Berechnungen. *Farbe* 14, 275-286.
- Valberg, A. (1971), An action spectrum for chromatic contrast and threshold processes. *Proc. 1st. European Biophysical Congress, Wien*.
- Weisenhorn, P. (1965), *Der Spektrale Farbintegrator und seine Entwicklung*. *Farbe* 14, 359-364.

DISCUSSION

Vos: I had difficulty in understanding what the meaning is of optimal colours in this context. One problem: did you really use optimal colours? Are not complementary colours only optimal if the spectrum is an equal energy spectrum? And then: why do you think optimal colours play an important role just in this context. I think they play a role in surface colours since they are of maximum reflectivity. However, I don't think they are in anyway special colours as to colour discrimination.

Valberg: According to the definition of an optimal colour (going back to E. Schrödinger), it is the object colour that for a certain purity and dominant wavelength has the highest possible luminance. What we have done is to simulate object colours by using a central field and a white surround, the central field being related to this surround by contrast (ger.: bezogene Farben). The stimuli projected in the central field obeyed the requirement necessary for optimal colours in that there are not more than two transitions from 0 to 1 in the relative spectral reflectance curves. The reason why we measured the discrimination ability of the eye to complementary optimal colour stimuli is an observation of rapid change of colour at the same positions within complementary "spectra". These "spectra" arise by projecting a slide equal to Fig. 1 of our paper through a prism. At each point within these two spectra, the relative spectral energy distribution is that of an optimal colour.

If you admit that wavelength discrimination is of relevance for colour theory, ideally, the same theory should be able to account for the fact that similar results are obtained with the complementary colour stimuli (unsaturated and light colours) plotted against the compensative wavelength scale.

Wyszecki: I must admit that I have some difficulty understanding the details of your study. Does the symmetry of color discrimination which you find for complementary pairs of colors, imply the existence of perfect circles around the white point in the ideal uniform color space?

Valberg: For a certain slitwidth, the loci of optimal colours are not related to equal distances from the white point in colour space. In fact, for such a series, saturation, lightness and hue varies continuously. In some regions of the spectra, the *JND* appears mainly as a lightness difference. In other regions, hue or saturation difference dominates. For pairs of complementary optimal colours however, the attributes combine in different ways to make the total *JND*.

Clarke: In your experiments no attempt was made to equalize the lightness. Do you think that you would have obtained a different kind of result if the observer had in fact tried to keep an equal lightness situation.

Valberg: This I do not know. Perhaps there would be symmetry if we equalize luminance. Operationally, equalization of luminance would at least make it easier to maintain complementarity of the stimuli for our procedure.

Bartleson: You have used a very restricted class of all possible complementaries, viz. spectral stimulus complements. This choice results in experimental evaluation of essentially the same stimulus *differences* in both cases if we assume that the discriminial mechanism involved is basic to the organism and not specific to the stimulus. Accordingly, I would think that you have provided an interesting alternative method for determining wavelength discrimination functions and, as such, the appearance of the stimuli are not really pertinent to the experiment. What, then, is the basis for inferring something about the perception of color from such results?

Valberg: We think that our results indicate a symmetry of discrimination about some indifferent or equilibrium state established by an adaptation stimulus. In our case this was the white Xe-light in the surround. To test this hypothesis, it would be interesting to repeat the experiments with a chromatic light source. Also we would like to do some experiments with a more general kind of complementarity not using optimal colour stimuli.

I agree that the appearance of the stimuli is not necessarily pertinent in discrimination experiments. At a second glance, however, the thresholds appear as combinations of different attributes. Although the differences of the components of the simple linear opponent-colours scheme applied in our case are equal for pairs of complementary stimuli, we know that the perceived differences of the component attributes cannot be constant because of the

5 DISCUSSION

non-linear relation between stimulus- and perceived magnitudes. Yet, the same stimulus difference gives rise to a *JND* for complementary pairs of stimuli and this might thus tell us something about the interplay of perceived attributes at threshold.

6 A ZONE-FLUCTUATION LINE ELEMENT DESCRIBING COLOUR DISCRIMINATION

J. J. VOS AND P. L. WALRAVEN

*Institute for Perception TNO,
Soesterberg, The Netherlands*

INTRODUCTION

Three ingredients for a theory of colour discrimination seem to be well established by now.

1. Colour vision is mediated by three receptor systems which we will denote by R , G , and B because of the approximate spectral location of their peak sensitivities. The actual course of the spectral sensitivities is still subject to debate, but the ideas seem to converge. Recently (Vos and Walraven, 1971) we have summarized and analyzed old and new material pertinent to the derivation of these spectral sensitivities from colorimetric properties of dichromats and arrived at the set of primaries shown in Fig. 1.

We will not discuss the arguments here and simply state that this set is compatible with the concept of dichromatism as trichromatism with one receptor system missing (Rushton, 1971, e.g.). The mathematical description of the three relative special sensitivities r_λ , g_λ and b_λ in terms of Judd's (1951) (x_λ, y_λ) specifications reads:

$$\begin{pmatrix} r_\lambda \\ g_\lambda \\ b_\lambda \end{pmatrix} = \begin{pmatrix} 0.19112 & 0.57903 & -0.03596 \\ -0.18576 & 0.42633 & 0.03060 \\ -0.00536 & -0.00536 & 0.00536 \end{pmatrix} \begin{pmatrix} x_\lambda/y_\lambda \\ 1 \\ 1/y_\lambda \end{pmatrix}$$

It should be noted that the curves of Fig. 1 give the spectral sensitivities of the three systems, not of the receptors individually. In order to fit phenomena like the Bezold-Brücke effect and the Weber behaviour of the three systems into the picture, we had to assume that the R , G , and B systems count different numbers of receptors:

$$N_R : N_G : N_B = 40 : 20 : 1$$

The sensitivity curves are, therefore, much more equivalent than apparent from Fig. 1, if plotted on a receptor basis; and even more so if we would correct them for intra-ocular absorption.

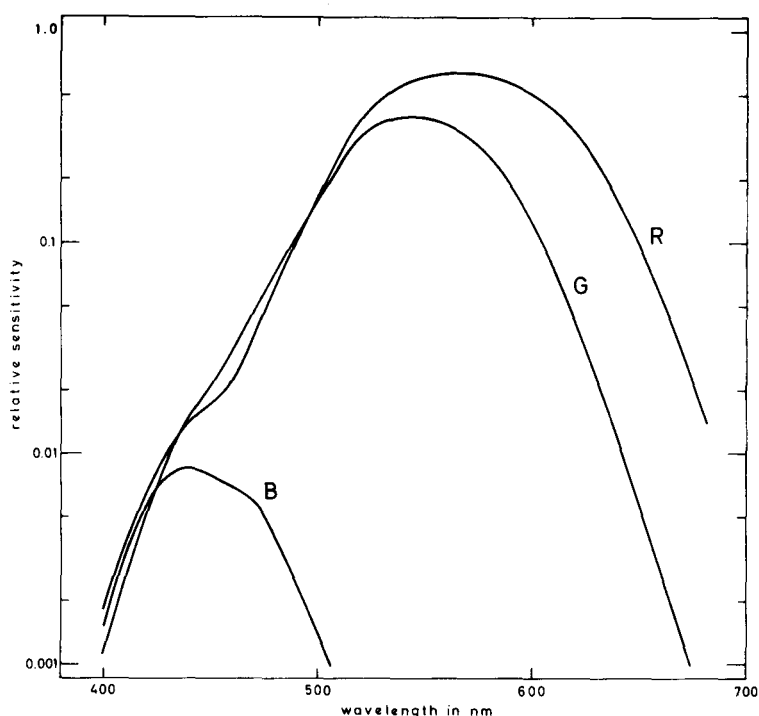


Fig. 1. The used set of spectral sensitivity functions for the foveal receptor systems (from Vos and Walraven, 1971).

2. Visual performance is essentially restricted by the intrinsic noise of the quantal message. Perhaps the most elegant way to illustrate that is by plotting the product of contrast sensitivity and spatial resolution against the light level. Van Meeteren and Vos (1972) showed that this product—as determined from the area underneath the spatial modulation-sensitivity functions—increases with the square root of the light intensity over as many as five decades (Fig. 2). This means that the visual system behaves in this respect as an ideal physical instrument and can be assumed to be essentially photon noise limited in its accuracy. As Walraven (1962) has shown, there is much reason to assume that the colour channels behave essentially similar. How the eye manages to handle its information so adequately over so many decades is a problem we are only beginning to understand (Bouman, 1969; Barlow, 1969; Rushton, 1969).

3. The three receptor signal is converted—already in the retina—into two antagonistic colour signals and one non-opponent brightness signal. The dis-

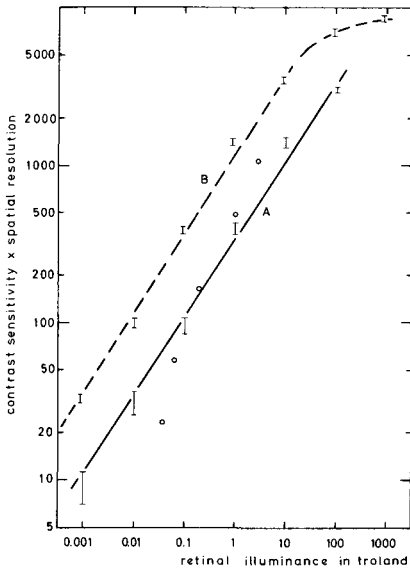


Fig. 2. The product of contrast sensitivity and spatial resolution—measure of the information transport capacity of the visual system—as a function of light intensity (from Van Meeteren and Vos, 1972).

cussions on this point have been hot in the past, but the storms have subsided in recent years (Jameson and Hurvich, 1968) after both the existence of spectrally distinctly different receptor systems (Marks, *et al.*, 1964) and the presence of antagonistic neural signals (DeValois, 1965) had been experimentally shown beyond reasonable doubt. The controversy seems now to be reduced to exactly which receptor systems feed into which conversion channels. Walraven (1962) has proposed a concrete zonal scheme of conversion. An updated version of it, as proposed by Vos and Walraven (1971) is shown in Fig. 3.

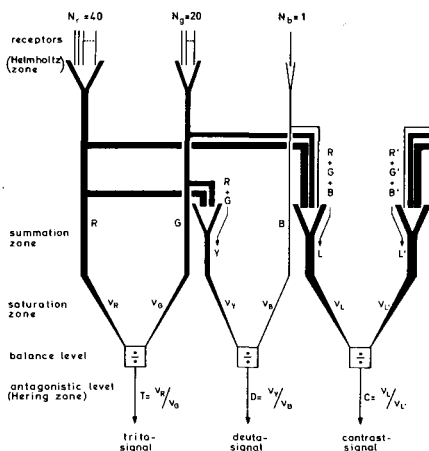


Fig. 3. Scheme of neural colour processing (according Vos and Walraven, 1971).

Abramov (1968) rejected the *G*-input to the *Y/B* channel, Guth (1968) the *B*-input to the *L*-channel as well. Wiesel and Hubel (1966), on the other hand, raised doubts as to the presence of a *G*-input into the *Y/B* system. The arguments for changes in the original Walraven/Bouman model are weak, however, since they are based on not too reliable data on the receptor primaries and on rather rough indications from electrophysiological experiments. We will, therefore, stick to the scheme of Fig. 3 which has some very distinct advantages: It recognizes the primacy of yellow; it is preferable above the other systems proposed as to mathematical elegance (see next section); and, not unrelated to the latter: it seems to be the most attractive choice from a phylogenetical point of view (Ladd-Franklin, 1909).

In the scheme we have indicated, by the narrowing shape of the channels, where we think signal transmission starts to be alinear due to adaptation processes. At low intensity levels, we need not bother about these effects of volume control, but at higher levels they turn out to play an important role. We then have to take v_R/v_G and v_Y/v_B , the ratios of the nerve signal intensities, instead of the ratio of the system stimulations *R/G* and *Y/B* as determinants of the colour response.

Based on these three concepts and on some less essential assumptions on the nature of adaptive processes, we have been able to describe wavelength and more general colour discrimination with fairly satisfactory accuracy. The *JND* λ vs. λ curves and the *JND* ellipses in the next are all calculated by computer programs, based upon our model. The detailed mathematics are submitted for publication (Vos and Walraven, 1972); the computer programs are available for those interested. Here we will try to explain the essential characteristics by visualizing the mathematics with an uniform chromaticity plot and to discuss the accuracy of the results obtained.

WAVELENGTH DISCRIMINATIONS AT LOW LIGHT LEVELS

Be *R*, *G*, and *B* the stimulation intensities of the three systems in numbers of quanta effectively absorbed. Then, each of them can be known with limited accuracy:

$$R \pm \sqrt{R}, G \pm \sqrt{G}, B \pm \sqrt{B}$$

Since these inaccuracies are independent of one another, each colour point in *R*, *G*, *B*-space is surrounded by an uncertainty ellipsoid with axes $\sqrt{R}$, $\sqrt{G}$, and $\sqrt{B}$ respectively. These ellipsoids determine, from place to place, and in any direction the *JND*-metric; i.e. they determine the distance which is significant on the basis of the intrinsic signal noise. It has advantages to contract the axes in such a way that the *JND*-ellipsoids become *JND*-spheres of constant size. The correct scale-contraction is obtained by a square root transformation

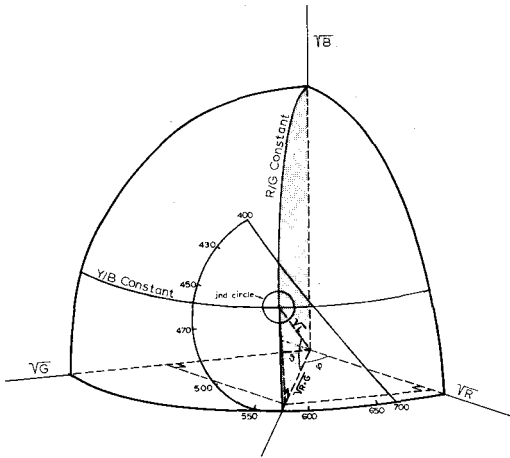


Fig. 4. The square-root diagrammatic representation of colour space.

$$\xi_R = \sqrt{R}, \xi_G = \sqrt{G}, \xi_B = \sqrt{B},$$

$$\text{since } \sigma_{\xi_X} = \frac{d\xi}{dX} \cdot \sigma_X = \frac{1}{2} \frac{1}{\sqrt{X}} \sqrt{X} = \frac{1}{2} = \text{constant indeed.}$$

In Fig. 4 this space contraction is applied.

Note that in this now euclidean space planes of constant brightness (spheres), of constant R/G (vertical planes through the B -axis), and of constant Y/B (vertical cones) are mutually orthogonal. This is what we had in mind when we hinted towards the mathematical elegance of the (R, G, B) to $(R/G, Y/B, R+G+B)$ transformation.

Since the contracted space is one of “uniform chromaticity”, we can measure colour discrimination simply by moving the unit JND -circle, as shown, over the sphere of constant luminance. The result will, apart from a factor, be independent of the light level, since the sphere does not change shape and only expands with the square root of luminance.

Algebraically it comes down to a line element of the shape

$$ds^2 = \frac{dR^2}{R} + \frac{dG^2}{G} + \frac{dB^2}{B}$$

which may be considered as closely related to the Schrödinger (1920) line element.

By moving the JND -circle along the spectral locus, we can find the course of wavelength discrimination over the spectrum. The results—but then obtained by the more accurate and more easy computerized calculus—is shown in Fig. 5 against the background of a set of more or less representative experimental results from literature.

The overall agreement seems to be quite satisfactory, but, actually, it is

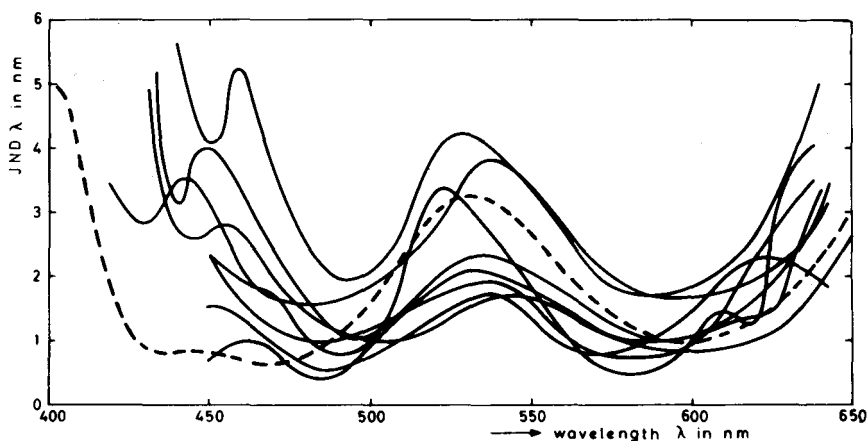


Fig. 5. Wavelength discrimination, calculated for low light levels, compared with experimental data.

-----: calculated relation; drawn curves: choice of curves from Wright and Pitt (1934) and from the compilation by Judd (1932).

difficult to say how satisfactory. In the first place, the same sort of accuracy is obtained—as we know—with other line elements. Apparently the choice is not very critical so that the fit does not prove at all the validity of the model. In the second place, to say that it roughly follows the experimental trend is in some way misleading. In particular, in the blue region the fit is not too good. Moreover, many of the curves in Fig. 5 show a distinct hump around 460 nm. We do not find an indication of such a hump in the theoretical curve.

An insight of what may be the cause of the discrepancy can be obtained from McCree's data on the change of wavelength discrimination with luminance (McCree, 1960).

It is clear that luminance does more than only determine the general level of colour discrimination. The hump at 460 nm, in particular, only appears at the higher light levels. We will have to investigate, therefore, what happens with increasing brightness.

WAVELENGTH DISCRIMINATION AT HIGH LIGHT LEVELS

At higher levels the nervous signal transmission becomes increasingly alinear. It has to become so, as otherwise adequate functioning of the visual system over as many as ten decades, roughly, would be impossible. Since a clear picture of how the mechanisms of volume control act is not available yet, we had to make a choice of a model on the basis of the principles of simplicity and adequacy. The model we used is, indeed, very simple but even so it can explain, apart from various features of colour discrimination: the shift in

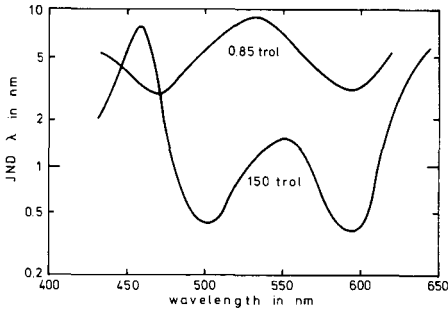


Fig. 6. Wavelength discrimination for a 75' field at two luminance levels (smoothed data from McCree, 1960).

colour appearance with the brightness level known as the Bezold-Brücke effect; the Weber-behaviour of colour discrimination at higher light levels; and the maintenance of the independency of the hue and brightness signals also at the Weber level.

The model is essentially that of a simple photon counter with refractory period. Such a counter starts to show saturation effects when the interval time becomes of the order of magnitude of the refractory period. At the same time the accuracy of signal transmission decreases, of course. Since the saturation effects were supposed to occur after the summation of R and G to Y (and of R , G , and B to L), we can only investigate the effects on colour discrimination by considering separately the R/G and Y/B channels.

An extended high brightness version of Fig. 4—in which the zonal conversion did not play a role yet—would not suffice. Consequently, the algebraic line element gets rather complex. Suffice it here to state that principally it has many characteristics in common with the Stiles' (1946) line element; and we refer to Vos and Walraven (1972) for the detailed mathematics. Here we may only mention, qualitatively, that the saturation model affects wavelength discrimination in two ways, essentially:

1. Because of the decreased accuracy of signal transmission, wavelength discrimination stops improving (Weber's law) and it will even start to break down beyond a certain intensity level.

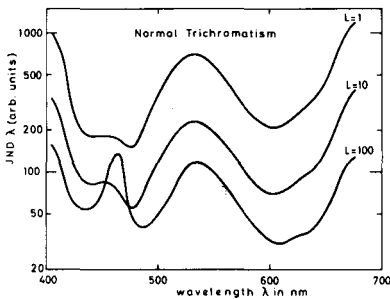


Fig. 7. Wavelength discrimination at three light levels, calculated on the basis of increasing saturation of the nervous channels.

2. Since the B-system has to carry the full brightness load with only a few receptors, it will be the most vulnerable system to these saturation and deterioration effects; and this, in particular, below 460 nm where it is equally sensitive to the other two systems and thus has to carry its full share. Therefore, these effects will show up first in the short wavelength region. The results of the exact calculations, as carried out by computer, are shown in Fig. 7. They, indeed, show the effect of deteriorating colour discrimination around 460 nm which we found to be manifest in the experimental data of McCree (Fig. 6).

JND ELLIPSES

Wavelength discrimination is the more sensitive method to test theoretical line elements; but at this conference emphasis is on discrimination between non-monochromatic pigment colours. We have derived, therefore, JND ellipses in the chromaticity diagram as well. In that case the line element takes a little bit more simple form:

$$ds_c^2 = f_{RG} [G(1 + G/G_0)dR + R(1 + R/R_0)dG]^2 + f_{YB} \{B(1 + B/B_0) + Y(1 + Y/Y_0)\}^2 (dR + dG)^2,$$

$$\text{with } f_{RG} = \frac{\frac{G}{1 + G/G_0 + G^3/G_1^3} + \frac{R}{1 + R/R_0 + R^3/R_1^3}}{\left[\frac{G(1 + G/G_0)}{1 + G/G_0 + G^3/G_1^3} + \frac{R(1 + R/R_0)}{1 + R/R_0 + R^3/R_1^3} \right]^2 \cdot (1 + R/R_0 + R^3/R_1^3) \cdot (1 + G/G_0 + G^3/G_1^3) RG}$$

$$f_{YB} = \frac{\frac{B}{1 + B/B_0 + B^3/B_1^3} + \frac{Y}{1 + Y/Y_0 + Y^3/Y_1^3}}{\left[\frac{B(1 + B/B_0)}{1 + B/B_0 + B^3/B_1^3} + \frac{Y(1 + Y/Y_0)}{1 + Y/Y_0 + Y^3/Y_1^3} \right]^2 \cdot (1 + Y/Y_0 + Y^3/Y_1^3) \cdot (1 + B/B_0 + B^3/B_1^3) YB}$$

and $Y = R + G.$

The constants R_0, G_0, B_0, Y_0 and R_1, G_1, B_1, Y_1 , represent levels of system overloading, and are related to the receptor population densities. On the basis $B_0 = 1$ they are $R_0 = 40, G_0 = 20, Y_0 = 20, R_1 = 80, G_1 = 40, B_1 = 2, Y_1 = 40$.

By computer program the ellipses can easily be calculated with this line element. Figs. 8, 9, and 10 show JND ellipses in the chromaticity diagram for three different light levels for some of the MacAdam points.

As might be expected, peculiarities like the reversal effect at 460 nm do not show up in these JND ellipses, as they are obscured to a large extent by the white veil in common to all colours. The influence of luminance remains restricted to smaller effects on shape and orientation.

The important question now is: How well do these ellipses describe the experimental data? To answer that question we need a more quantitative measure of

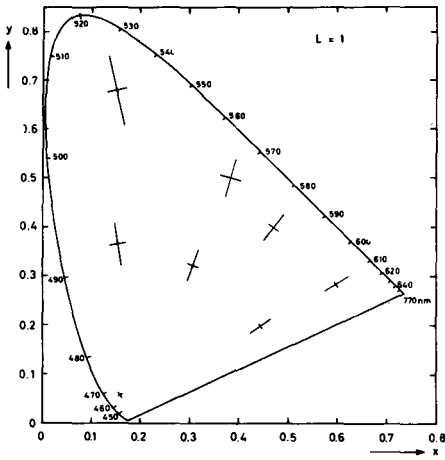


Fig. 8. *JND* ellipses calculated for low luminance levels.

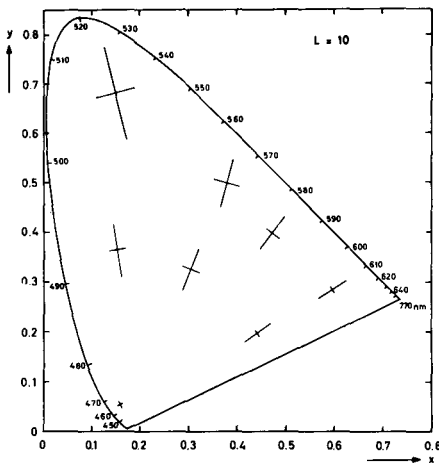


Fig. 9. *JND* ellipses calculated for medium luminance levels.

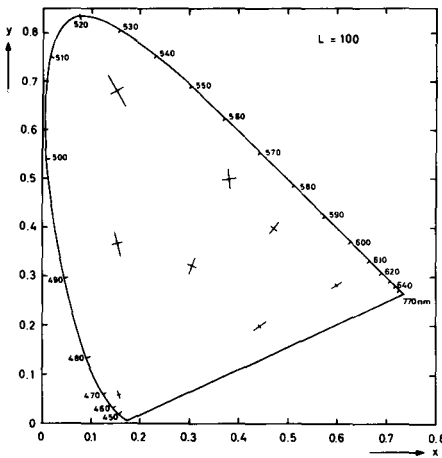


Fig. 10. *JND* ellipses calculated for high luminance levels.

similarity than that of subjective appraisal. We, therefore, developed a "dissimilarity index" D between two sets of ellipses that takes into account, in a weighted manner, the inequality of corresponding ellipses in the sets as to size, elongation, and orientation. The exact calculation can be found in Vos and Walraven (1972). $D=0$ means a complete fit *on the average*. It does not mean that the two sets are identical since, for instance, a clockwise difference in orientation between two ellipses may just be compensated by a counterclockwise difference between two others. Therefore, we give, apart from D itself, the accuracy of D based on the spread in dissimilarity of the constituent pairs of ellipses. It shows in some way the significance of D .

Let us first compare our theoretical results with the experimental ellipses by MacAdam (1942):

MacAdam vs. theor.	$L = 1$	$D = 0.16 \pm 0.13$
"	"	"
"	$L = 10$	$D = 0.16 \pm 0.12$
"	"	"
"	$L = 30$	$D = 0.19 \pm 0.12$
"	"	"
"	$L = 100$	$D = 0.31 \pm 0.13$

It will be clear that the experimental fit is about equally good for the three lowest brightness levels, but that it begins to decrease beyond $L = 30$.

Since we have no standard of how good $D = 0.16$ is, we may compare our fit with those obtained by Stiles—as taken from Wyszecki and Stiles (1967)—, and by Chickering (1967) by optimizing the Friele-MacAdam formula:

MacAdam vs. Stiles	$D = 0.38 \pm 0.10$
Mac Adam vs. Chickering	$D = 0.01 \pm 0.08$

Stiles' fit is definitely worse than that obtained with our line element, due to the plumpness of his ellipses. Chickering's fit is almost exact for the average ellipse—as it should be after his optimization procedures. However, this precision seems to be somewhat absurd in view of the standard deviation.

This is illustrated in another way by the following comparison between sets of experimental ellipses obtained by Brown and MacAdam (1949)—as taken from Wyszecki and Stiles (1967):

obs. WRJB vs. obs. DLM	$D = 0.38 \pm 0.07$
obs. DLM vs. obs. PGN	$D = 0.58 \pm 0.08$
obs. PGN vs. obs. WRJB	$D = 0.28 \pm 0.08$

Apparently differences between the results of even experienced observers are much larger than those between our calculated set of ellipses and the experimental set of MacAdam. For that reason we may question the sense of an optimization procedure to get close fit to the results of only one observer. As long as we stay with the description of the MacAdam ellipses—ellipsoids may be another question, but we have not yet analyzed these in terms of the zone-

fluctuation line element—the simple description in terms of the uncomplicated low light level line element seems to be adequate as the present state of lack of experimental knowledge.

REFERENCES

- Abramov, I. (1968), Further analysis of the responses of IGN cells. *J. Opt. Soc. Amer.* 58, 574-579.
- Barlow, H. B. (1969), Coding of light intensity by the cat retina. In: *Processing of optical data by organisms and by machines*. pp. 384-396, (Ed. W. Reichardt) Academic Press, New York.
- Bouman, M. A. (1969), My image of the retina. *Qu. Rev. Biophys.* 2, 25-64.
- Brown, W. R. J. and MacAdam, D. L. (1949), Visual sensitivities to combined chromaticity and luminance differences. *J. Opt. Soc. Amer.* 39, 808-834.
- Chickering, K. D. (1967), Optimization of the MacAdam-modified Friele Color-difference formula. *J. Opt. Soc. Amer.* 57, 537-541.
- DeValois, R. L. (1965), Analysis and coding of color vision in the primate visual system. *Proc. Cold Spring Harbor Symp. on Quant. Biol.* 30, 567-579.
- Guth, S. L., Alexander, J. V., Chumby, J. I., Gillman, C. B., and Patterson, M. M., Factors affecting luminance additivity at threshold among normal and color blind-subjects and elaborations of a trichromatic opponent-colors theory. *Vision Res.* 8, 913-928.
- Jameson, D. and Hurvich, L. M. (1968), Opponent-response functions related to measured cone photopigments. *J. Opt. Soc. Amer.* 58, 429-430.
- Judd, D. B. (1932), Chromatic sensitivity to stimulus differences. *J. Opt. Soc. Amer.* 22, 72-108.
- Judd, D. B. (1951), Secr. Rep. "Colorimetry and artificial daylight". In: *Proc. 12th Session CIE, Stockholm Vol I, Techn. Comm. No. 7*, p. 11.
- Ladd-Franklin, C. (1909), The nature of colour sensations. Add. to *Helmholtz Treatise on Physiol. Optics*, 3rd Ed., Dover Publ., 1962, pp. 662.
- MacAdam, D. L. (1942), Visual sensitivities to color differences in daylight. *J. Opt. Soc. Amer.* 32, 247-274.
- McCree, K. J. (1960), Small field tritanopia and the effects of voluntary fixation. *Optica Acta* 7, 317-323.
- Marks, W. B., Dobelle, W. H., and MacNicol Jr, E. F. (1964), Visual pigments of single primate cones. *Science* 143, 1181-1183.
- Pitt, F. H. G. (1944), The nature of normal trichromatic and dichromatic vision, *Proc. Roy. Soc. Lond. B* 132, 101-117.
- Rushton, W. A. H. (1969), Light and dark adaptation in the retina. In: *Processing of optical data by organisms and by machines*. pp. 544-564, (Ed. W. Reichardt) Academic Press, New York.
- Rushton, W. A. H. (1971), Color vision: An approach through the cone pigments. *Invest. Ophthalm.* 10, 311-322.
- Schrödinger, E. (1920), Grundlinien einer Theorie der Farbenmetrik im Tagessehen. *Ann. Physik* 63, 397-456; 481-520.
- Stiles, W. S. (1946), A modified Helmholtz line element in brightness-colour space. *Proc. Phys. Soc. Lond.* 58, 41-65.
- Van Meeteren, A. and Vos, J. J. (1972), Resolution and contrast sensitivity at low luminances. *Vision res.* 12, 825-33.
- Vos, J. J. and Walraven, P. L. (1971), On the derivation of the foveal receptor primaries. *Vision Res.* 11, 799-818.
- Vos, J. J. and Walraven, P. L. (1972), An analytical description of the line element in the zone-fluctuation model of colour vision I and II. *Vision Res.* 12, in press.
- Walraven, P. L. (1962), On the mechanisms of colour vision. Thesis Utrecht, Inst. for Perception Ed.

6 DISCUSSION

- Wiesel, T. N. and Hubel D. (1966), Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *J. Neurophysiol.* 29, 1115-1156.
- Wright, W. D. and Pitt, F. H. G. (1934), Hue discrimination in normal color vision. *Proc. Phys. Soc.* 46, 459.
- Wyszecki, T. and Stiles W. S. (1967), *Color science*, Tables 6.12 and 6.15, John Wiley, New York.

DISCUSSION

Wright: In the wavelength discrimination curve, the subsidiary minimum in the curve at about 440 nm corresponds to the more critical discrimination associated with the change in hue from blue to violet. The violet hue itself is associated with the subsidiary hump in the spectral sensitivity curve of the red receptor process. The agreement between theory and experiment in Dr. Vos' analysis will therefore be critically dependent on the sensitivity functions used by the authors (Fig. 1, Vos and Walraven) and its poor agreement in the blue and violet might be eliminated by a relatively small change in the shape of the *R* curve in the 440 nm region of the spectrum.

Vos: Of course one can get the right JND λ versus λ curve by changing the set of primaries—but we had not that freedom, since our primaries were determined independently from the colour mixture laws of di- and trichromats. The apparent poor fit in the blue largely disappears, however, if we take into account the overloading effect of the *B*-system described. From the relative failure of the *Y/B* system at higher light levels one can understand both the occurrence of the subsidiary minimum at 440 nm (Fig. 7) and the appearance of violet as a hue—the latter being associated with the then relative dominance in the “far blue” of the *R/G* system at high light levels.

Wright: I don't think I agree with you. You can understand these things already without any zonal processing.

Vos: You are right, in so far, that the overloading of the *B*-system already occurs at a pre-zonal stage. And so we do not need the zone model for this particular purpose.

MacAdam: How do you account for the negative curvature of the equiluminous surface that is known to be required to represent observed colour discrimination ellipses and Wright's colour discrimination segments for equiluminous colours?

Vos: I have not much thought in terms of positive or negative curvature, to be honest. However, does not your comment mean that in your opinion, my set of predicted ellipses cannot possibly describe your set of experimental ellipses adequately?

Wyszecki: I believe the smaller ellipse at the “white point” is, at least in part, a result of the “crispening” effect. Discrimination increases when the test color is close to that of the surround, conditions like they have been used in MacAdam's 1942 work.

Vos: Coming back to Dr. MacAdam's question—may the following be of relevance? Fig. 4 is essentially the low luminance uniform colour plot. At high light levels, to keep the plot of uniform spacing, the scales along the axes gradually transform from square root to logarithmic. And so does additive luminance ($V_\lambda = R_\lambda + G_\lambda + B_\lambda$) change to multiplicative ($V_\lambda = R^{0.67} G^{0.32} B^{0.01}$). The latter surface has negative curvature. The fit with the MacAdam ellipses is slightly worse, however, than for those calculated for low luminances.

Stiles: It is realized that you have presented here a much simplified version of your model and that in the more sophisticated model all sorts of cross-terms creep in due to the zonal processing—and as a result the question of the curvature may quite well turn out to be not so controversial.

Vos: That implies that our measure for discrepancy is not sophisticated enough?

Stiles: It may be that the more sophisticated model gives a better distribution of discrepancies.

May I put another question:

In the fuller theory of your line-element, shortly to be published, do you rely, in fixing any constants, on the precise stimulus level at which the relation between the noise threshold + the stimulus intensity changes from the square root form of de Vries-Rose, to a form which approximates, finally, to the constant Weber fraction law?

Vos: These saturation levels—or may I better use the term overloading levels since saturation has other meanings, as you did correctly point out—play an essential role in our derivation of the primaries. I agree, that you cannot easily indicate the transition from the square root region to the Weber region in the experimental threshold curves—masked as it may be by adaptative effects like the shrinking of integration time and area with increasing light level.

7 A NEW COLOR MODEL

SHERMAN LEON (LEE) GUTH

*Department of Psychology, Indiana University
Bloomington, Indiana 47401*

It has been shown that foveal thresholds and brightness matches cannot be predicted on the basis of a scalar additivity rule applied to the monochromatic components of lights. That is, Abney's law is not valid for these tasks. On the other hand, flicker photometry is known to be an additive system in the sense that flicker-photometric matches of polychromatic lights can be precisely predicted on the basis of the scalar sum of the luminances of their monochromatic components.

We (1969) have used an opponent-colors theory to explain why Abney's law holds for flicker-photometric judgments but not for detection or direct brightness matching judgments. Since flicker-photometric matches are made only after chromatic fusion has occurred, flicker photometry is assumed to tap only a nonopponent system. However, detection and direct brightness matching are mediated by opponent-colors mechanisms *and* by the nonopponent mechanism. Failures of Abney's law are assumed to be due to cancellations or other interactions which occur within chromatic mechanisms; therefore, they are found with detection and brightness-matching tasks, but not with flicker photometry. The theory also implies that the foveal spectral sensitivity function obtained by flicker photometry should differ from those obtained by the other two methods. Very recent studies from three different laboratories [Guth and Lodge (1971), Kaiser (1971), Boynton and Wagner (1971)] have confirmed that the direct matching and threshold spectral sensitivity functions show increased sensitivities in both the long and short wavelength ranges. Also, according to Kaiser (1971), Wagner's recent data contradict accepted conclusions about step-by-step sensitivity functions in that the functions deviate from flicker in the same way as direct brightness matching.

These facts, together with the finding (Guth, 1969) that vector addition is a useful alternative to Abney's law point to a conceptual revision of our current colorimetric and photometric models.

The model to be presented is basically an opponent-colors theory with the usual nonopponent achromatic (*A*) mechanism, a red vs. green (*T*) mechanism, and a blue vs. yellow (*D*) mechanism. (Actually, we are not convinced that the unitary hues are red, green, yellow and blue, or even that the nonopponent system signals an achromatic sensation. We have also come to realize that hue reports from even a theoretically ideal unilateral dichromat do not necessarily identify the unique hues. Nevertheless, we shall continue to use the convenient traditional color names.) The model is still in its preliminary stage, but it is expected that it will ultimately account for as many facts of color vision as have other opponent-colors models. However, it has the unique property of being based upon a simple three-dimensional Euclidean space where the brightness or detectability of a stimulus is represented by the length of origin-bound color vectors in that space. The three dimensions of the space reflect contributions to the color percept from the nonopponent as well as the red vs. green and blue vs. yellow mechanisms. At threshold levels, our model takes its simplest form. However, as brightness is increased, the model is complicated mainly by an increase in chromatic components and by the emergence of a hypothesized inhibitory effect of the blue vs. yellow system on the output of the red vs. green system.

Within the model, vector loadings on the achromatic dimension will relate to the CIE V_λ function, whereas the total origin-bound vector lengths will be related to spectral sensitivities obtained by direct matching or foveal detection procedures. The relative saturation of an equal brightness spectrum will be predicted on the basis of the ratio of the sum of the chromatic loadings to the sum of all loadings for each unit length vector. Wavelength discrimination between equal brightness spectral lights will be related to the Euclidean distance between neighboring unit-length vectors.

The unique predictions to be made by the model will concern heterochromatic additivity failures. The model will predict red-plus-green additivity failures at both threshold and high-brightness levels. Blue plus yellowish-green mixtures will be subadditive within the threshold model but superadditive within the high-brightness model. It will also be shown that Wyszecki's equal-brightness contours for equal-luminance stimuli are predicted by the model. Finally, we will show that the high-brightness model predicts curvilinear constant-hue loci within the CIE chromaticity diagram.

Since important changes take place in color processing as intensity level is varied, we shall treat threshold and high-brightness models separately. The threshold model is not intended to account for rod intrusions, but rather assumes that retinal stimulation is limited to the rod-free portion of the fovea.

1. THE THRESHOLD MODEL

The A , T , D threshold model is defined in terms of the CIE $\bar{x}$, $\bar{y}$ and z distribution coefficients by the following transformation equations:

$$\begin{aligned}\bar{a} &= .000\bar{x} + .954\bar{y} + .010z \\ \bar{i} &= .799\bar{x} - .646\bar{y} - .167z \\ \bar{d} &= .000\bar{x} - .058\bar{y} + .030z\end{aligned}\quad (1)$$

Fig. 1 shows a plot of the $\bar{a}$, $\bar{i}$ and $\bar{d}$ values.

The chromatic $\bar{a}$ function defines a relative luminous efficiency function which is very similar to the CIE V_λ function. However, it has been made to have somewhat increased sensitivity in the blue end of the spectrum as does Judd's (1951) correction. The symbol which we use to identify spectral luminous efficiencies within the model is V_λ^* . We shall refer to V_λ^* as *achromatic luminance* to distinguish it from *vector luminance* which is defined below. Obviously, if one transforms the X , Y , Z tristimulus values for any stimulus into A , T , D values according to the transformation equations given above, then the resulting A value is a luminance term comparable to Y .

The chromatic $\bar{i}$ and $\bar{d}$ functions identify the responses of the red vs. green and blue vs. yellow opponent systems, respectively. Positive $\bar{i}$ values imply a redish signal, whereas positive $\bar{d}$ values imply a blueish signal. Note that the contribution from the blue-yellow system is very small relative to the contribution of the red-green system. This is consistent with many theories of color vision which recognize that there is relatively little blue-yellow encoding at near-threshold levels. The wavelength at which the red-green system changes from red to green is between 570 and 575 nm, whereas the corresponding crosspoint for the blue-yellow system is at about 492 nm. These crosspoints should be interpreted only as preliminary values, and a minor rotation of the axes of the space would place the crosspoints closer to 500 and 575 as demanded by some color models.

The unique property of the threshold model is the manner in which it combines the outputs of the achromatic and chromatic systems. It is assumed that the achromatic system, the blue vs. yellow system, and the red vs. green system represent orthogonal axes in a three-dimensional Euclidean space. The total length of any vector in this space is called its *vector luminance*, abbreviated L^{**} . Vectors of equal length are assumed to represent lights of equal brightness, or, in the case of the threshold model, lights of equal detectability. The length of each equal-radiance spectral vector of wavelength λ is defined in the Euclidean space as $(\bar{a}_\lambda^2 + \bar{i}_\lambda^2 + \bar{d}_\lambda^2)^{1/2}$. V_λ^{**} is a special L^{**} quantity which defines the relative vector-luminance of equal-radiance spectral stimuli. It is of immediate interest to compare vector luminance, V_λ^{**} , with the CIE V_λ function and with $\bar{a}$ (or V_λ^*) which represents the relative spectral sensitivity of the achro-

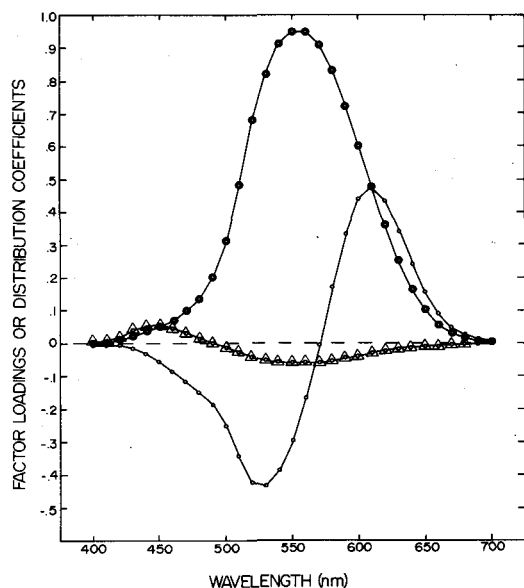


Fig. 1. Coordinates of unit-radiance spectral vectors on the A (large circles), T (small circles) and D (triangles) axes of the threshold vector model.

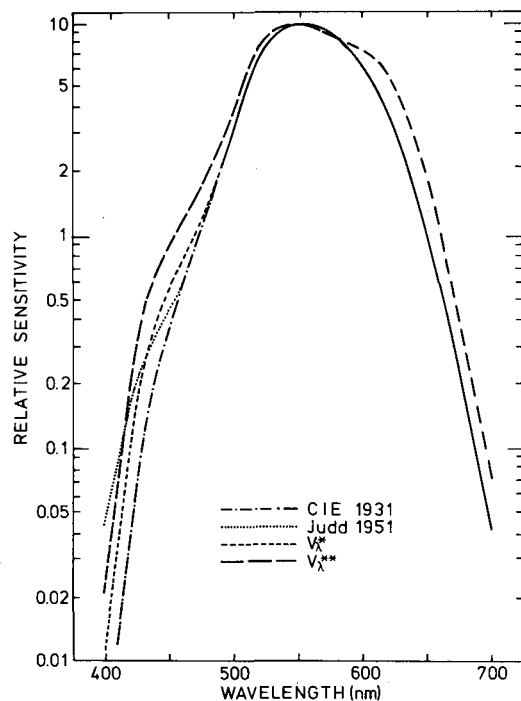


Fig. 2. Relative luminous efficiencies V_λ according to *CIE* and to Judd, and on the basis of the definition of the achromatic luminance(*) and vector luminance(**) in the threshold vector model.

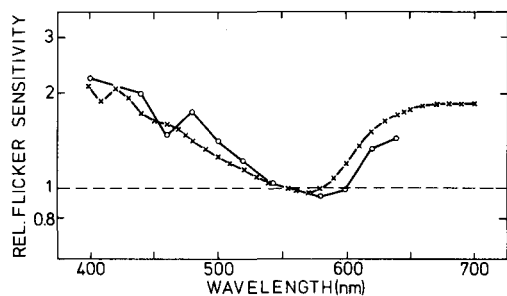


Fig. 3. Mean foveal spectral sensitivities from direct brightness matching divided by flicker-photometric sensitivities (open circles) and V_λ^{**} divided by V_λ^{*} ("x's") from the threshold vector model. Normalization is at 555 nm.

matic system. The CIE V_λ function (·-·- in Fig. 2) is the function showing the lowest sensitivity in the blue end of the spectrum. The function showing the highest sensitivity at 400 nm (. . . .) represent Judd's (1951) correction of the CIE values below about 460 nm. The $\bar{a}$ or V_λ^* function (- - - -) is the function which falls between the CIE V_λ function and Judd's correction in the blue end. At wavelengths longer than about 490 nm, the CIE V_λ function and V_λ^* are essentially identical. The function which defines vector luminance (- - -) shows an increase in sensitivity in both the red and blue ends of the spectrum. As mentioned earlier there is much evidence that spectral sensitivities as evaluated by threshold or direct brightness matching procedures show increased sensitivity in the red and blue ends of the spectrum when compared to flicker-photometric data. Our interest in this question led us to complete an experiment in which each of six naïve subjects were tested using both threshold and flicker-photometric procedures in the same apparatus. A 45' circular field was used, and the 191 *td* white standard for direct brightness matching was presented in the right half of the field. The two halves of the field were separated with a fine black line. Each observer made a total of four observations per wavelength per day on at least four different days. The curve drawn with open circles in Fig. 3 shows the mean threshold spectral sensitivities relative to flicker-photometric sensitivities (That is, the ratio of threshold to flicker sensitivities). Normalization is at 555 nm. Note the increased sensitivity in the red and blue ends of the spectrum shown by the matching data. Each of the six subjects showed similar increases in the spectrum extremes. For comparison, the curve drawn with "x's" shows V_λ^{**} relative to V_λ^* with normalization at 555 nm. Note that the experimental differences are well matched by the predictions made by the vector model.

Since we have defined the length of each equal-radiance vector as V_λ^{**} , it is a simple matter to transform the equal-radiance coordinates to equal-luminance coordinates. All we need do is to divide $\bar{a}$, $\bar{i}$ and $\bar{d}$ for each wavelength by its V_λ^{**} value. The resulting values are called A_L , T_L and D_L . A plot of these coordinates, which can be thought of as tristimulus values for an equal-luminance spectrum, is shown in Fig. 4. Note that in the yellow and nearby portion of the spectrum, the achromatic system has by far the dominant response. This reflects the fact that the spectrum is desaturated in that region. On the other hand, the contribution from the chromatic systems are relatively high in the end portions of the spectrum. Within the model, a straightforward measure of the relative saturation of the equal brightness spectral stimuli can be obtained by dividing the chromatic response, $(T_L + D_L)$ by the sum of all responses $(A_L + T_L + D_L)$. The predicted relative saturation of the spectrum at threshold levels is shown in Fig. 5.

Within the model, the distances between equal-length vectors should be related to wavelength discrimination for equal-brightness stimuli. To predict

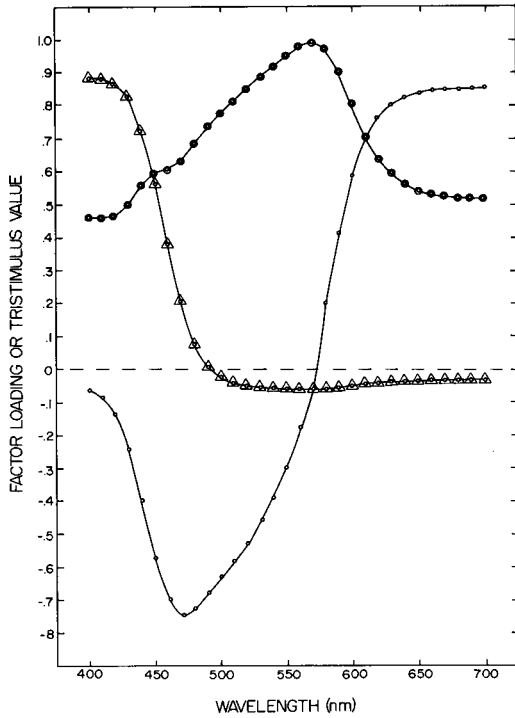


Fig. 4. Coordinates of unit vector luminance spectral vectors on the A (large circles), T (small circles) and D (triangles) axes of the threshold vector model.

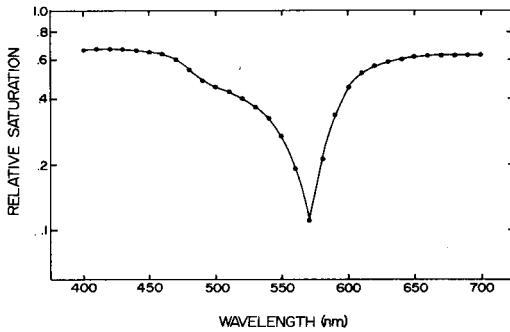


Fig. 5. Relative saturation of the equal vector luminance spectrum according to the A , T , D , threshold vector model.

wavelength discrimination, we should define a line element within the space and determine the $\Delta\lambda$ which corresponds to that distance at all points in the spectrum. We have not as yet completed that analysis, but an approximate idea of what that function would look like is shown in Fig. 6. The points in that figure were derived by computing the Euclidean distances between unit-length vectors at 10 nm intervals. The distances above and below each wavelength in 10 nm steps were then averaged and the reciprocal of that distance was plotted. The function suggests that low-level wavelength discrimination is

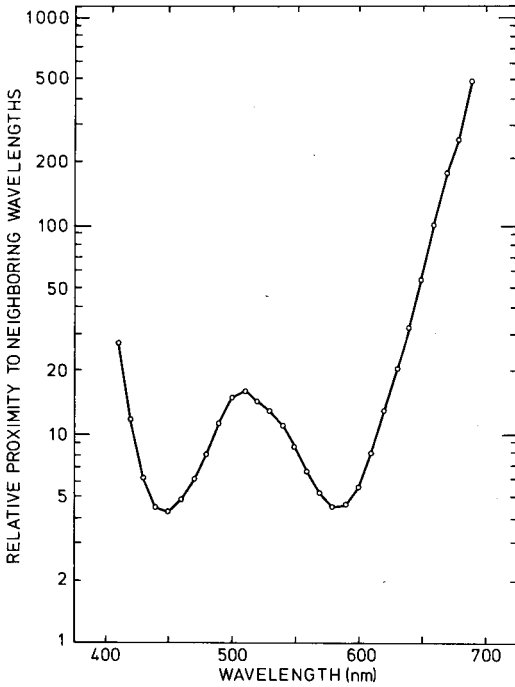


Fig. 6. Approximate wavelength discrimination for the equal vector luminance according to the A , T , D , threshold vector model.

at best at about 450 and 590 nm, and the curve is strikingly similar to experimental data obtained by Weale (1951).

Fig. 7 shows the projections of chromatic stimuli into the equal achromatic plane. That is, it is a plot of T_A and D_A which are obtained by dividing T_L and D_L by A_L —a procedure which projects the spectral vectors into the unit A plane. Within our model, the distances between pairs of wavelengths in this diagram represent distances to be expected between chromatic stimuli if their achromatic components (i.e., their achromatic luminances) are equal. This would correspond closely to an experiment in which the investigator controlled for the traditional luminance quantity, but more interestingly, it corresponds

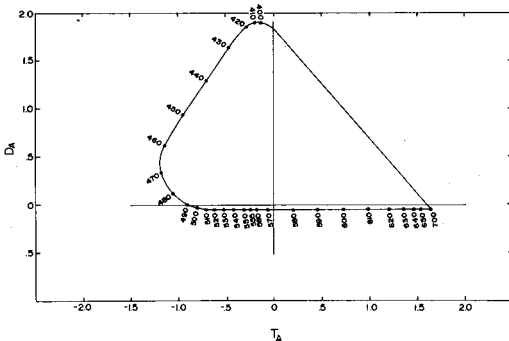


Fig. 7. The unit achromatic response plane of the A , T , D threshold vector model.

to the experiment reported in this symposium by Boynton and Wagner. They attempted to scale the differences between chromatic stimuli after equating their achromatic components using the border minimization technique. To the extent that border minimization in fact equates stimuli for their achromatic components, then we predict that distances between wavelengths shown in Fig. 7 will correspond to the distances which they would obtain in a threshold-level experiment.

The most unique property of our color vision model is that it will predict the heterochromatic additivity failures which we found using foveally viewed threshold-level stimuli. The predictions proceed in a straightforward manner.

A convenient measure of heterochromatic additivity failures between two wavelengths of magnitudes L_i^{**} and L_j^{**} is:

$$\cos \theta_{i,j} = \frac{(L_r^{**})^2 - (L_i^{**})^2 - (L_j^{**})^2}{2 L_i^{**} L_j^{**}}, \quad (2)$$

where L_r^{**} is the magnitude of the resultant obtained from mixing L_i^{**} and L_j^{**} . A cosine of plus one implies complete additivity, whereas a cosine of minus one implies complete subtraction. In our experiments we typically determine the values of L_i^{**} and L_j^{**} which, when mixed, yield a unit vector-luminance (i.e., threshold) resultant.

The solid lines of Fig. 8 show average cosines obtained in our laboratory over a period of several years. The values shown are very close to those given in Table 2 of our 1969 paper, but data from a few more subjects have been added. Each function shows cosines which describe additivity failures between the wavelength indicated by the upright arrow and the wavelengths shown on the abscissa. It is a simple matter to predict corresponding cosines from the A , T , D threshold space. For wavelengths i and j , which are each of unit length,

$$\cos \theta_{i,j} = (A_{Li})(A_{Lj}) + (T_{Li})(T_{Lj}) + (D_{Li})(D_{Lj}) \quad (3)$$

The predicted cosines are shown with dashes in Fig. 8. The excellent agreement between the obtained and predicted functions is not surprising since we have already shown (Guth *et al.*, 1969) that a vector model based upon color-mixture data predicts threshold-level additivity failures.

It is important to note, however, that we have now made the vector space a simple transformation of the CIE X , Y , Z system. Therefore, we suggest that vector luminance, L^{**} , should be used to specify the detectabilities of foveally viewed signal lights. This merely requires the use of the transformation equations (1) to derive A , T and D from tristimulus values X , Y and Z , and the subsequent computation of:

$$L^{**} = (A^2 + T^2 + D^2)^{1/2}. \quad (4)$$

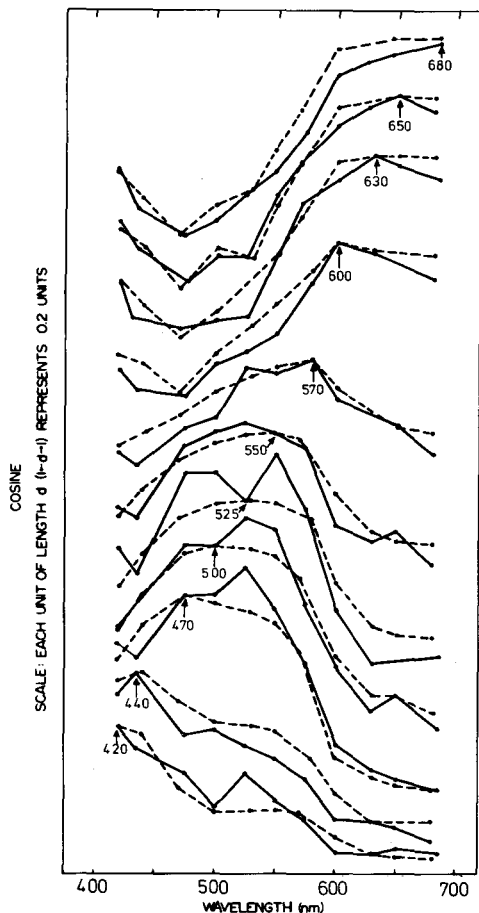


Fig. 8. Each function shows obtained (solid lines) and predicted (dashes) cosines between the wavelength indicated with the upright arrow and the wavelengths given on the abscissa. The cosine at each arrow is +1.0.

The functions shown in Fig. 8 demonstrate that vector luminance will predict the detectabilities of bichromatic mixtures. Data presented by Lodge, Huff and Guth (1969) imply that vector addition will also predict the detectabilities of trichromatic (and, via Grassmann's laws, polychromatic) mixtures. Therefore our A , T , D threshold space allows a new and useful photometry for near-threshold lights.

II. THE HIGH-BRIGHTNESS MODEL

The following transformation equations define the A , T , D high-brightness vector space, which is in a very preliminary stage.

$$\begin{aligned}\bar{a} &= 0.000\bar{x} + 0.893\bar{y} + 0.009z \\ \bar{t} &= 1.070\bar{x} - 0.864\bar{y} - 0.223z \\ \bar{d} &= 0.000\bar{x} - 0.078\bar{y} + 0.040z\end{aligned}\tag{5}$$

The only difference between these equations and those given earlier for the threshold model is that the contribution from the achromatic system has been considerably reduced. Therefore, the $\bar{a}$ coefficients have been reduced by a constant. To renormalize the space, the coefficients assigned to the $\bar{i}$ and $\bar{d}$ values have been increased by another constant. The related functions for an equal-radiance space are shown with solid lines in Fig. 9.

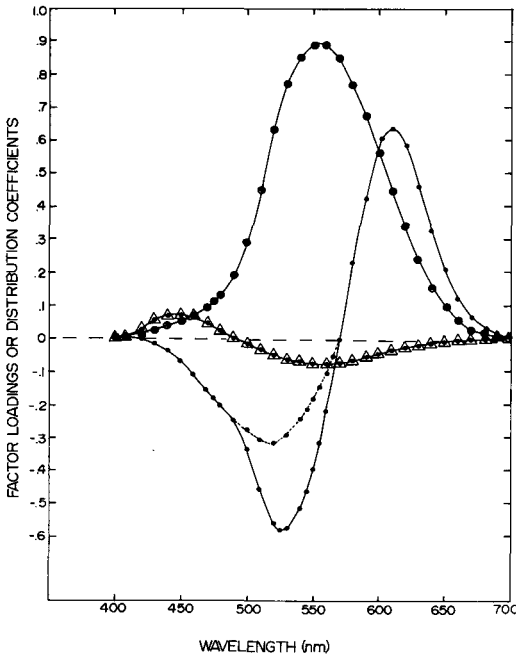


Fig. 9. Coordinates of unit radiance spectral vectors on the A (large circles), T (small circles) and D (triangles) axes of the high-brightness vector model. The dotted lines show T values after they are inhibited by D .

However, as mentioned earlier, we hypothesize that the blue vs. yellow system inhibits the red vs. green system. The magnitude of inhibition is given by the quantity $\frac{1}{1+c(d)}$ where d is the activity of the blue vs. yellow system and c is assigned the value of 15.0 for this particular model. This is a popular formulation for neural inhibition. It has the properties that the inhibition factor is unity when the quantity doing the inhibiting is zero, and that as the intensity of the stimulus doing the inhibiting gets higher and higher, the inhibitory factor becomes directly proportional to intensity. For the present high-brightness model, the nature of the inhibitory activity is that yellow activity inhibits green activity and that blue activity inhibits red activity. (This rule will probably be modified as the model develops.) That is, in terms of the signs which we have assigned to chromatic activity, $+\bar{d}$ inhibits $+\bar{i}$ and $-\bar{d}$ inhibits $-\bar{i}$. The dashes in Fig. 9 show the result of that inhibition. The existence of the inhibitory effect makes the prediction of a relative spectral efficiency function based upon total vector length (i.e., V_{λ}^{**} values) somewhat more difficult than it was for

the threshold model. What we need to know is the relative radiance that a subject would require for each wavelength in order to yield unit-length vectors in the A, T, D space. However, as intensity is raised, inhibition from $-\bar{d}$ to $-\bar{i}$ increases. This is a negative feedback loop, and to determine the factor by which one must multiply the $\bar{a}$, $\bar{i}$ and $\bar{d}$ values in order to yield a unit length vector, a 4th degree equation must be solved. Fortunately, this is no problem for our computers. If we divide the $\bar{a}$, $\bar{i}$ and $\bar{d}$ values for a particular wavelength by the quantity obtained for V_{λ}^{**} , then we obtain the values A_L , T_L and D_L . For those portions of the spectrum where inhibitory interactions do not occur, $A_L^2 + T_L^2 + D_L^2 = 1.0$. However, when inhibition occurs, $A_L^2 + T_L^2 + D_L^2 > 1.0$. Only after T_L is inhibited by D_L do we obtain a new T_{Li} quantity which, when squared and added to A_L^2 and D_L^2 yields a quantity which again sums to unity. Of immediate interest is the shape of the V_{λ}^{**} function because it should resemble relative luminance efficiency functions obtained by the direct brightness matching method. At this stage of the model, the V_{λ}^{**} function indicates the expected increased sensitivity in the spectral extremes when compared to the achromatic luminosity function, but the increase is too great to provide a good fit to experimental data. Many of the high-brightness models which we have examined do predict direct matching sensitivities, but those models have lacked other desirable properties. We are confident that the final model will incorporate an acceptable V_{λ}^{**} function as well as the other characteristics which will be described below.

As was true at threshold levels, the chromatic components of unit-length vectors divided by the sum of all components is an obvious measure of the saturation of an equal-brightness spectrum. Relative saturations (S) can be computed for wavelengths where inhibition does (S_i) or does not (S) occur according to

$$S = \frac{T_L + D_L}{A_L + T_L + D_L} \text{ or } S_i = \frac{T_{Li} + D_L}{A_L + T_{Li} + D_L}.$$

Fig. 10 shows such ratios. The function obviously has many characteristics in common with experimental measures of the saturation of the spectrum.

Fig. 11, which shows the inverse of distances between neighboring equal-length vectors, was derived in the same way as was Fig. 6, except that inhibited values were used when appropriate. The dashes indicate interpolation within the small ranges where inhibitory interactions suddenly occur between chromatic systems. The curve is especially interesting because of the emergence of the new minimum point at around 500 nm which is the region where one would expect to find good wavelength discrimination for an equal-brightness, high level spectrum.

The inhibitory mechanism built into the high-brightness model allows extremely interesting predictions regarding suprathreshold heterochromatic

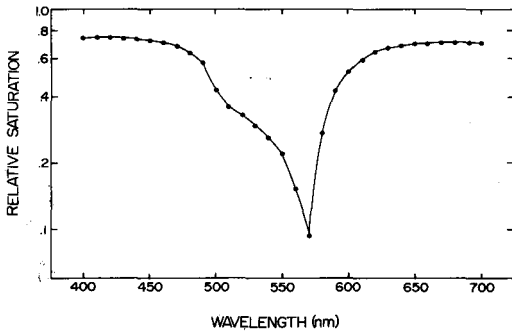


Fig. 10. Relative saturation (S or S_r) of the equal-brightness spectrum according to the A , T , D high brightness vector model.

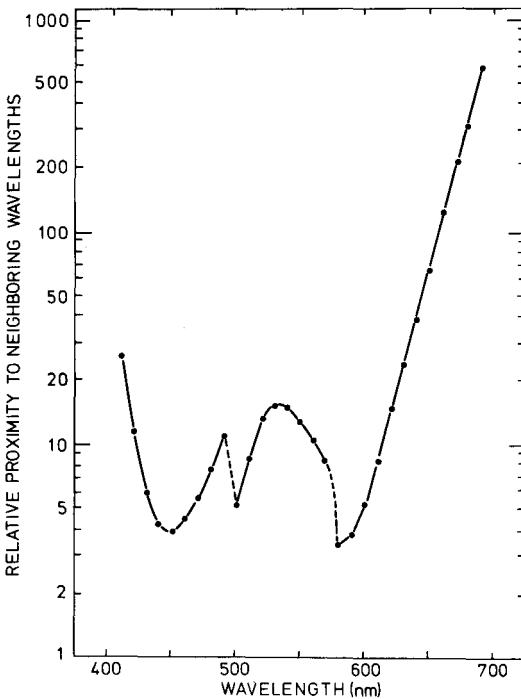


Fig. 11. Approximate wavelength discrimination for the equal-brightness spectrum according to the A , T , D high brightness vector model.

additivity failures. We have shown (Guth, 1969) that mixtures of various wavelengths with a red light are less bright than is predicted by a simple additivity principle. This is especially true for green-plus-red mixtures which are even more subadditive at high levels than they are at threshold. On the other hand, some blues and greens are superadditive at photopic levels. These facts are summarized with solid lines in Fig. 12. The data of Fig. 12 are from an experiment reported in our 1969 paper. Included in that experiment were conditions in which we combined equal portions of various monochromatic stimuli into many bichromatic mixtures and determined the amount of each compo-

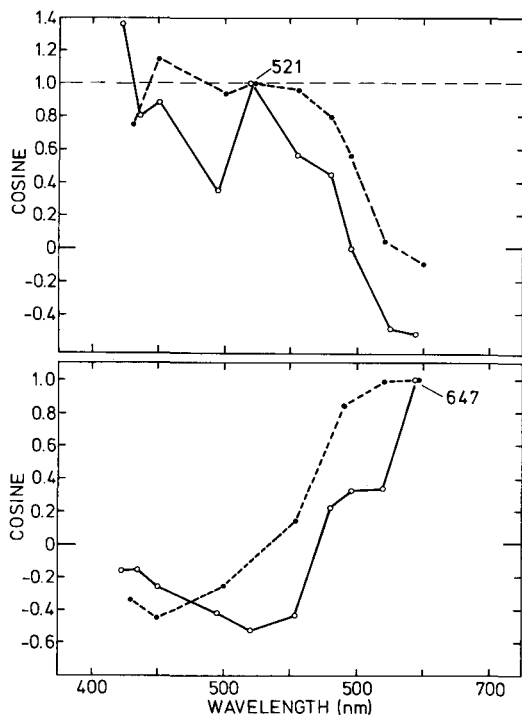


Fig. 12. Each function shows obtained (solid lines) and predicted (dashes) cosines between the wavelength indicated with the arrow and the wavelength marked on the abscissa.

ment in the mixture when the mixture was adjusted to a unit-brightness level by the subject. We then used Eq. (2) to derive the cosine which accounted for the data. Obtained cosines when spectral stimuli were mixed with 521 or 647 nm are shown in the figure. Of special interest is the 521 nm function. When reds were added to 521 nm, the usual large subadditive effects obtained and the resulting cosines are highly negative values. When 500 nm was combined with 521 nm, subadditive effects also were observed, but when 420 nm was used, an enhancement effect occurred. (The cosines of greater than unity which describe enhancement have no meaning in terms of real angles, but they are nevertheless useful in describing data.)

In order to predict the suprathreshold additivity data, we asked of the model what would happen if we added, say, 0.50 vector luminance units of 521 nm to 0.50 units of 450 nm. Since inhibitory interactions are assumed to occur after opponent processes, combining stimuli requires addition of the uninhibited A_L , T_L , and D_L values. When we add the D values for the 450 nm stimulus to the D values for the 521 nm stimulus, the result is a positive D value. Therefore, the D component resulting from the mixture of 521 and 450 nm does not inhibit the T component, the latter of which is relatively high because we added in the uninhibited T quantity for 521 nm as demanded by the model. *The disinhibition of T causes an enhancement effect when 450 nm is added to 521 nm,*

and the dashed line shown in Fig. 12 goes above unity at 450 nm. Notice that there is also a slight dip in the function at 500 nm indicating subadditive effects at that wavelength and that when long wavelengths are added to 521 nm, there are severe subadditive effects. The bottom portion of Fig. 12 shows in dashes the predicted cosine between the wavelength shown on the abscissa and 647 nm according to the high-brightness vector model. Fig. 12 in and of itself justifies the presentation of our preliminary high-brightness model, for the ability of the model to account for both enhancement and subadditive effects in heterochromatic additivity at high levels is surprising.

There exist other data in the literature concerning heterochromatic additivity failures. Wyszecki (1967) has obtained brightness judgments for many equal-luminance color samples specified within the x, y chromaticity diagram. To summarize the results of the experiment, loci of constant brightness-to-luminance ratios were drawn. Fig. 13 shows one of these diagrams. To derive predicted brightness-to-luminance ratios from our model, we merely transformed many x, y, z coordinates into A, T, D values using transformation Eqs (2). The resulting vectors were made to have equal achromatic luminance values by dividing A, T and D by A . Then D was made to inhibit T whenever appropriate, and the length of the resulting vector was computed as $L^{**} = (1.0^2 + T^2 + D^2)^{1/2}$. The model's approximation of a brightness-to-luminance ratio for a unit achromatic luminance vector is then $L^{**}/1.0$ (i.e., vector luminance over achromatic luminance). Loci of constant vector luminance-to-achromatic luminance ratios are shown in the x, y chromaticity diagram in Fig. 14. There seems to be no similarity between this figure and Fig. 13, but the difference is illusory. If we draw smooth loci using only those x, y points which are in the range defined by Wyszecki's experiment, we obtain Fig. 15, which is indeed similar to Fig. 13.

Another interesting characteristic of the high-brightness model is that it predicts curvilinear loci of constant hue in the x, y diagram. Within the model, it is assumed that the dominant hue of a color percept is related to the D -to- T ratio. To show these loci of constant hue, we computed the D -to- T ratio for each of the points shown in Fig. 14. We used the inhibited T value whenever appropriate. Fig. 16 shows the approximate loci.

In conclusion, we should like to offer our own guess as to the properties which the model will have in its final form. The major difference will be that yellow responses will inhibit red as well as green responses. Blue responses might also inhibit green as well as red. With such a rule, the model will show curvilinear constant-hue loci in additional areas of the chromaticity diagram. We hope that it will also account for MacAdam's ellipses, the Bezold-Brücke effect and spectral sensitivities as obtained by direct brightness matching.

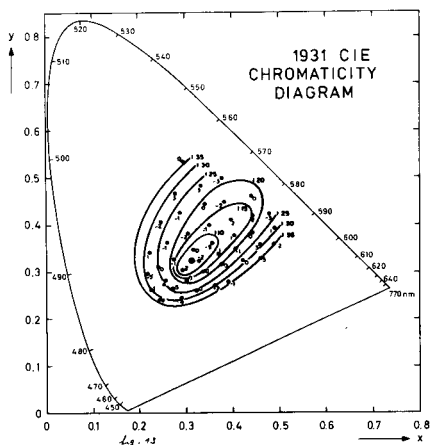


Fig. 13. Loci of constant brightness/luminance ratios from Wyszecki (1967).

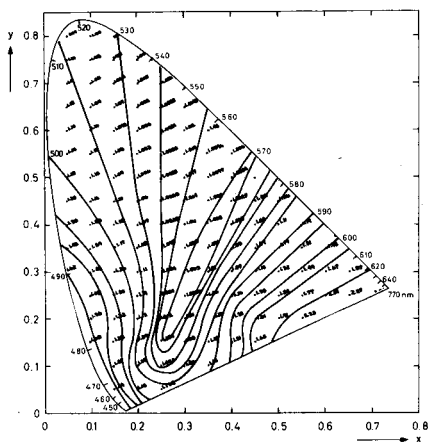


Fig. 14. Loci of constant brightness/luminance ratios predicted by the high-brightness A, T, D vector model.

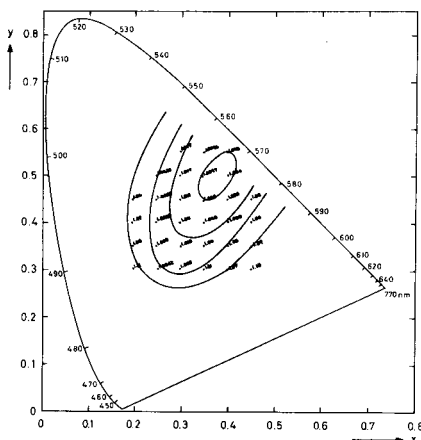


Fig. 15. Same as Fig. 14, except that loci are drawn through a subset of x, y points.

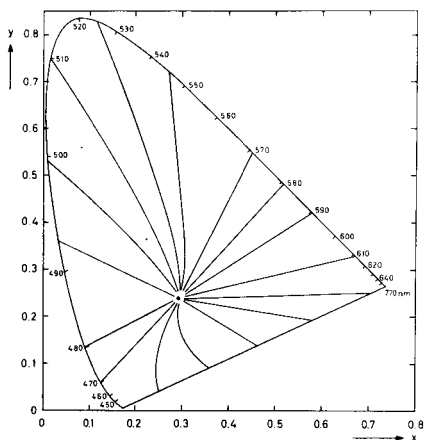


Fig. 16. Loci of constant hue according to the high-brightness A, T, D vector model.

REFERENCES

- Boynton, R. M. and Wagner, H. G. (Personal Communication).
 Guth, S. L., Donley, N. J., and Marrocco, R. T. (1969). On luminance additivity and related topics. *Vision Res.*, 9, 537-575.
 Guth, S. L. and Lodge, H. R. (1971). Tests of a theory regarding spectral sensitivities of normals and dichromats. Paper presented at Midwest Psychological Association, Detroit.
 Judd, D. B. (1951). Report of U. S. Secretariat Committee on Colorimetry and Artificial Daylight, Proc. CIE Stockholm Vol. 1, part 7, p. 11.

- Kaiser, P. K. (1971). Minimally distinct border as a preferred psychophysical criterion in visual heterochromatic photometry, *J. Opt. Soc. Amer.* 61, 966-971.
- Lodge, H. R., Huff, J. E., and Guth, S. L. (1969). Test of a vector model for near-threshold luminance additivity. *J. Opt. Soc. Amer.* 59, 1533. (Abstr.)
- Weale, R. A. (1951). Hue-discrimination in para-central parts of the human retina measured at different luminance levels. *J. Physiol.*, 113, 115-123.
- Wyszecki, G. (1967). Correlate for lightness in terms of CIE chromaticity coordinates and luminous reflectance. *J. Opt. Soc. Amer.* 57, 254-257.

DISCUSSION

Nimeroff: In Fig. 16, why do the constant hue lines converge on the chromaticity coordinate point (0.29, 0.24)? Other investigations have chosen white-points nearer to (0.33, 0.33).

Guth: Through the influence of Dr. Vos, I have come to believe that it is not a good idea to build a model on the assumption that zero output from chromatic systems implies a "white" perception. For all we know, the non-opponent system signals a percept which looks somewhat blue. It is better to build a model which accounts for other facts of color vision, and then infer what the non-opponent system signals. To my knowledge, there is absolutely no set of experiments which allow us to know what percept the non-opponent system conveys. It would be nice if we would have a unilateral achromat.

Vos: My argument in favour of a neutral point — let us drop the word white — more towards the blue corner is that there we expect the only really invariable hue with respect to the Bezold-Brücke effect. It lies on the cross point of the line of constant R/G , connecting the experimentally found 570 nm invariable hue point with the tritanopic confusion center, and the line of constant Y/B through the experimentally found invariable hue points in the blue (476 nm) and in the purple and the point $x = 1$ on the alychne. As can be seen from the figure, though, this neutral point lies quite a bit more off white than even Dr. Guth's neutral point.

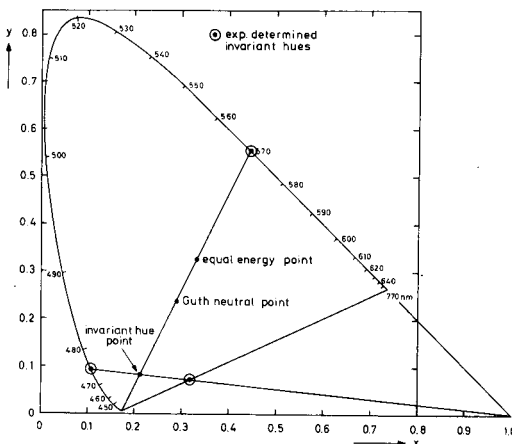


Fig. 17. The construction of the invariant hue point.

MacAdam: How do you account for differences between Figs. 14 and 15 of your paper?

Guth: There is no difference. It only shows that, on the basis of a limited subset of data, you easily draw lines like Dr. Wyszecki drew them.

Hunt: In Eq. 1 on p. 84 were the size of the constants in the equation for $\bar{a}$, (and $\bar{d}$) chosen to optimize the predictions of the model, or were they based on some other data?

Guth: They were chosen to optimize many things, not for specific features like invariant hues or pure colors. They were obtained by how they would predict luminance additivity, saturation, and the difference between the sensitivity of the achromatic system as opposed to the total vector length. And on top of that I had to look at transformations which also had the inhibition of the Y/B system on the R/G system. It is not optimal yet, based as it is on trial and error iteration.

Hunt: Fig. 7 is reminiscent of uniform chromaticity diagrams; have you explored the model's ability to depict colour differences?

Guth: Fig. 7 holds for threshold. We tried, of course, to draw circles and transform then to the xy diagram. The results were terrible. You get almost straight lines. However, by introducing inhibition of the R/G system by the Y/B system at higher light levels, you get nice ellipses. The preliminary results are encouraging and I think there will lie the answer.

MacAdam: Please be specific about your superadditivity effect. What wavelengths did you use? What was the observer's response? Specifically, with values of luminance or radiance.

Guth: Let us give an example of the experiment, that may make things more clear. If you match 647 nm against white and then, for instance 420 nm against the same white, then half the amount 647 nm + half the amount 420 nm do not match the white. They look darker. That is what I call sub-additivity. If you do the same with 520 nm + 420 nm, then you get upper-additivity.

Clarke: I think that what everybody is puzzled about are the values of the cosine in Fig. 12, they do not exceed unity except on the violet region. I wonder whether you could explain this a bit more.

Guth: The cosine is not a theoretical interpretation, it is just a measure of additivity. Cosine=1 means vectors in opposite directions (180°), i.e. subtraction. Most of the data in Fig. 12 show minor or major degrees of subadditivity. However, violet + green turns out to be super additive.

8 BRIGHTNESS UNITS AND THRESHOLD UNITS

B. H. CRAWFORD*

*Imperial College,
London, Great Britain*

The investigation of just perceptible chromaticity shifts by direct methods has led to several unexpected results which will be reported in full elsewhere, but one chance discovery is, perhaps, of wider interest and significance. It concerns the difference between the visual effectiveness of radiation as judged by brightness and by liminal increment. It has often been assumed that visual effectiveness will be the same by both methods and in many cases, indeed, it is, at least approximately. In the case to be described, however, the difference may be startlingly large, to the extent of being obvious by merely qualitative observation.

The principle of the experiment is quite simple. A field of view of 2° square can be illuminated in two ways: either a white in one half can be compared with illumination of the other half by light confined to a narrow spectral band, or the whole field is white (the same white) and the coloured light is added to half of it until a colour difference is just perceptible. Suitable calibration of the apparatus enables an observer to compare a determination of brightness equality with one of liminal brightness increment. Nothing is changed in the apparatus between the two determinations except by the insertion or removal of a simple opaque screen over half the field of view.

If the narrow spectral band is in the deep blue part of the spectrum, centred at 430 nm, then, as would be expected, observers differ considerably in their assessment of the blue field by brightness comparison with the white field (of energy distribution approximately that of CIE source D65). Differences in macular and lenticular pigmentation are presumably the prime cause of these differences between observers, nearly 100 : 1 for those so far tested. The unexpected thing is that when the observers determine the liminal increase of 430 nm

* Due to circumstances, the author could not attend the meeting. The paper was not orally presented at the symposium.

which can just be detected as a colour difference when added to half the white field, they differ but little between themselves.

The numerical results for three observers are shown in Table I and, graphically, in Fig. 1.

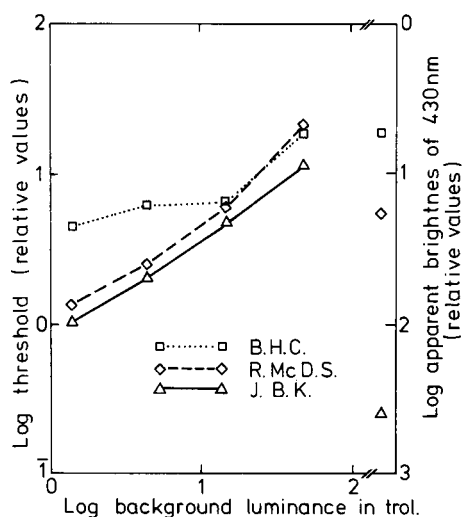


Fig. 1. Relation between threshold for light of 430 nm and background luminance (white, *D65*) for three observers; also their relative assessments of brightness of 430 nm. Note large discrepancies between the threshold and brightness assessments of light of 430 nm.

TABLE I Relative brightness and relative threshold of a 430 nm stimulus

Observer	Log relative brightness by direct comparison	Log relative threshold on background of	
		1.4tr	48tr
B.H.C.	0.72	0.65	1.26
R. Mc D.S.	1.26	0.13	1.32
J.B.K.	2.61	0.02	0.15

In order to be sure that no errors of instrumental origin were coming in, observers inspected the field appearances set by each other and made closely consecutive check settings. As mentioned above, differences between observers were qualitatively obvious even without recording quantitative measurements.

In discussing these curious results, it must first be noted that direct comparison of the brightnesses of the two halves of a field when there is a large colour difference between them may not give the same result as a flicker comparison, which is generally accepted as giving the correct measure of luminance. The divergences are, however, small in comparison with the results now being discussed, although they have interesting features which will be treated in another context in some future publication. As a present theory, it would seem that at moderately high levels of field luminance the threshold for light from the violet part of the spectrum is determined by the effect of the field on the blue visual

mechanism only, even though the field contains radiation through the whole spectrum. This supposition is strongly supported by the following experiment.

The white field in the apparatus was produced as the recombination of a continuous spectrum (primarily so that it could be tailored to the distribution of *D65*). It was therefore a simple matter to remove any desired part of the spectrum by insertion of an opaque screen in the appropriate position. The result found was that removal of the red, red plus yellow, or red plus yellow plus green part of the spectrum had no significant effect on the measurement of threshold for 430 nm made on the remainder of the spectrum recombined to form the main field. Fig. 2 shows an example of actual measurements. Thresh-

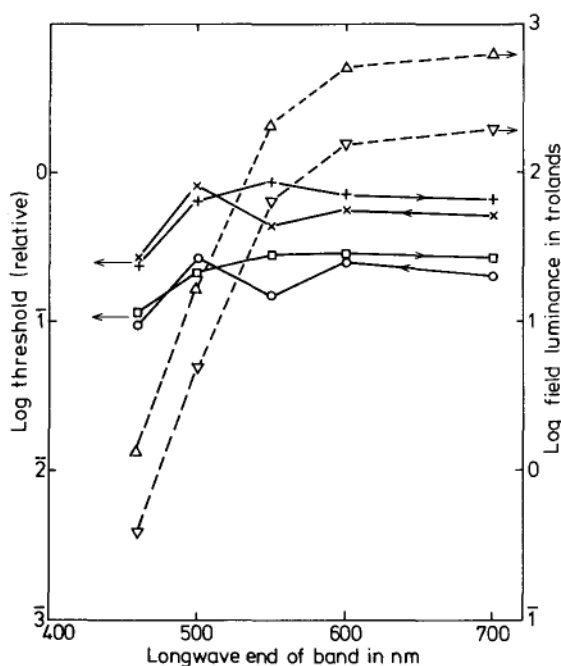


Fig. 2. Full lines: relation between threshold for light of 430 nm and composition of background (series of bands from 400 nm to plotted wavelengths). Broken lines: relation between background luminance and composition (same series of bands).

old measurements were made at progressive stages of removal of spectrum starting from the red end, also in the reverse order to minimise any effects of adaptation or drift of observer's criterion. Measurements were also made of the luminance of the main field at the various stages. The remnant of spectrum can be reduced to the band 400 to 500 nm without significant effect on the measured threshold for 430 nm, although the luminance has dropped by a factor of nearly 40 to 1 (from 630 tr to 17 tr, or from 200 tr to 4.9 tr in the two cases shown in Fig. 2).

The way in which threshold perception of 430 nm responds only to the wave-band 400 to 500 nm and ignores the rest of the spectrum from 500 to 760 nm is remarkable. There is even a hint in the measurements that the longer wavelength parts of the spectrum may have a slightly inhibiting effect, but this must be pursued in more appropriately designed experiments before significance can be attached to the hint.

In relation to colour metrics and the idea of a "uniform chromaticity space", these present results show that the latter concept can never be universally applied, although it may possibly have some practical use as an approximation within strictly limited conditions of observation. The transformation of just perceptible colour differences into the notation of chromaticity in a system such as that of the CIE will always introduce artifacts into the appearance of the results which may be misleading, especially as their existence is easily overlooked. Fig. 3 shows the sort of thing which can happen as a result of such transformation, showing differences between observers which are largely fictitious.

In view of the "Helmholtz-memorial" theme of the Driebergen conference, it may be of interest to note that the apparatus used for the work described above is an adaptation of the principles of the Helmholtz colorimeter.

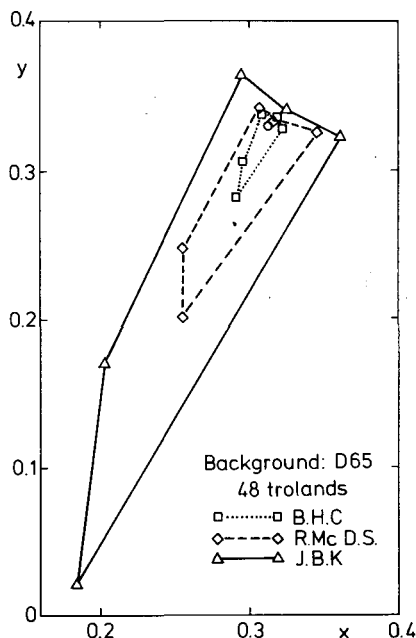


Fig. 3. Just perceptible colour differences plotted in relation to the individual brightness assessments shown in Fig. 1 and the incremental thresholds for wavelengths 430, 470, 520, 580 and 650 nm measured on white background of 48 trolands. The incremental thresholds are transformed arithmetically into chromaticity shifts by use of the standard colorimetric tables.

9 THE INFLUENCE OF THE FUNDAMENTAL PRIMARIES ON CHROMATIC ADAPTATION AND COLOUR-DIFFERENCE EVALUATION UNDER DIFFERENT ILLUMINANTS

H. TERSTIEGE

*Bundesanstalt für Materialprüfung, 1
Berlin 45, U. d. Eichen 87*

Since König one has tried to determine the fundamental system of the human eye. In this fundamental system it is possible to describe processes of colour perception, as long as they are concerned to retinal base and not yet transformed to opponent signals, by means of a simple diagonal matrix.

Determinations of spectral tristimulus values however result in colour-matching functions of a real primary system R, G, B , which is connected with the fundamental primary system P, D, T by a transformation matrix K of 9 constants:

$$\begin{pmatrix} P \\ D \\ T \end{pmatrix} = \begin{pmatrix} k_1 & k_2 & k_3 \\ k_4 & k_5 & k_6 \\ k_7 & k_8 & k_9 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

$$\begin{pmatrix} P \\ D \\ T \end{pmatrix} = K \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

The evaluation of this transformation matrix or the fundamental colour-matching functions has been tried in different ways:

1. from the data of the achromatic points of the dichromats which can be obtained from the intersection of the confusion lines (König, Judd . . .)
2. from the results of threshold investigations (Stiles, Wald . . .)
3. from photometric investigations of the retina (Rushton, Marks, Dobelle, MacNichols . . .)

Regarding Fig. 1, one can see the tremendous spread of the chromaticity coordinates of the fundamental primaries determined by different authors in that way. While the fundamental primaries P and T of the various authors are

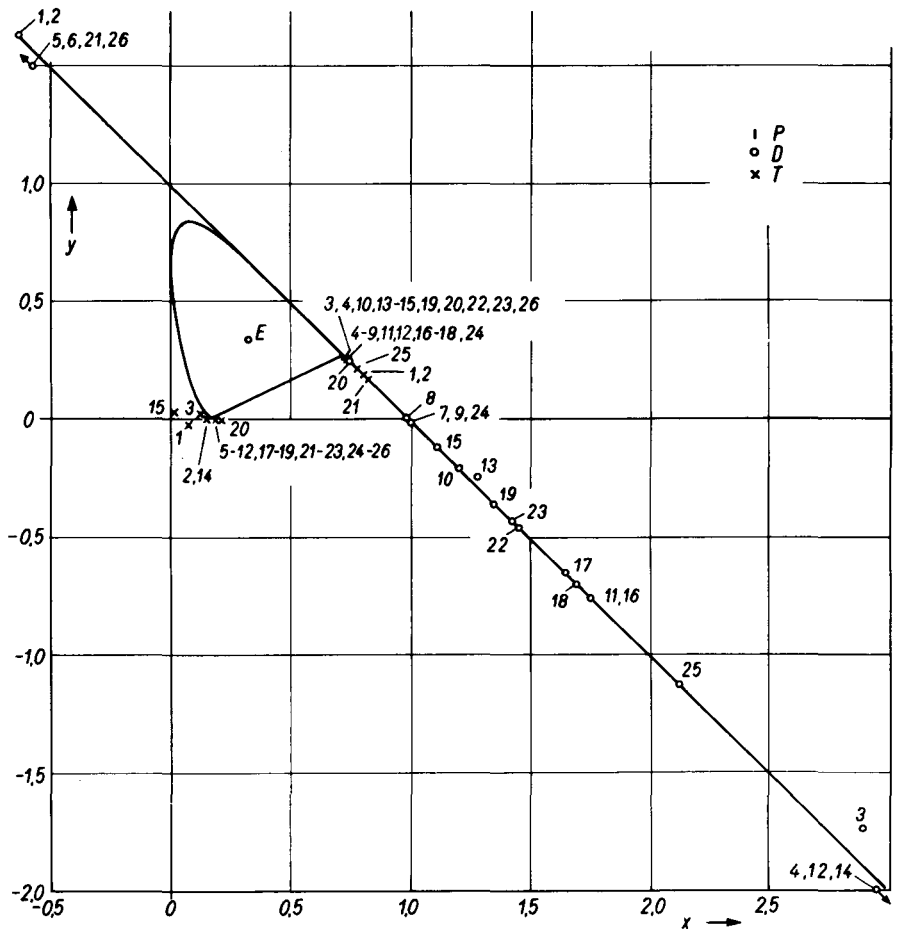


Fig. 1. Colour points of fundamental primaries from different authors:

1. König-Judd } Judd	11. MacAdam (1956)	20. Burnham-Brewer
2. König-Ives } (1941)	12. MacAdam (1957)	(Br. 1954)
3. Wright (1934)	13. Pitt (1944)	21. Fedorov <i>et al.</i> (1953)
4. Walters (1942)	14. Pitt (1944)	22. Szekeres (1948)
5. Thomson-Wright (1953)	15. Bouma (1942)	23. Kustarev (1965)
6. Sperling (1960)	16. Justova (1948)	24. Boynton (1969)
7. Pitt-Judd } Judd,	17. Justova (1950)	25. Stilés (1946)
8. Judd } (1944, 45)	18. Nyberg-Justova (1957)	26. De Vries (1946)
9. LeGrand (1948)	19. Lobanova-Rautian	
10. Steffen (1955/56)	(1952)	

fairly close together the primary D spreads in the chromaticity diagram along the prolongation to both sides of the spectrum locus.

According to various authors therefore fundamental primaries $P^* D^* T^*$ exist, which are connected with the not yet determined true fundamental primaries $P D T$ as follows:

$$\begin{pmatrix} P^* \\ D^* \\ T^* \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \begin{pmatrix} P \\ D \\ T \end{pmatrix}$$

$$\begin{pmatrix} P^* \\ D^* \\ T^* \end{pmatrix} = M \begin{pmatrix} P \\ D \\ T \end{pmatrix}$$

The matrix M affects correspondingly the fundamental colour-matching curves $\bar{p}(\lambda)$, $\bar{d}(\lambda)$, $\bar{i}(\lambda)$ and from this result different fundamental colour-matching functions $\bar{p}^*(\lambda)$, $\bar{d}^*(\lambda)$, $\bar{i}^*(\lambda)$:

$$\begin{pmatrix} \bar{p}^*(\lambda) \\ \bar{d}^*(\lambda) \\ \bar{i}^*(\lambda) \end{pmatrix} = (M^T)^{-1} \begin{pmatrix} \bar{p}(\lambda) \\ \bar{d}(\lambda) \\ \bar{i}(\lambda) \end{pmatrix}$$

In this formula $(M^T)^{-1}$ is the transposed inverse of the matrix M .

If we limit ourselves to the case of the most uncertain fundamental primary D then P and T may be regarded as well known ($a_1=1$; $a_2=a_3=0$; $c_1=c_2=0$; $c_3=1$), we get the fundamental colour-matching functions in dependence of the variation of the fundamental primary D :

$$\begin{pmatrix} \bar{p}^*(\lambda) \\ \bar{d}^*(\lambda) \\ \bar{i}^*(\lambda) \end{pmatrix} = \begin{pmatrix} 1 - \frac{b_1}{b_2} & 0 \\ 0 & \frac{1}{b_2} \\ 0 - \frac{b_3}{b_2} & 1 \end{pmatrix} \begin{pmatrix} \bar{p}(\lambda) \\ \bar{d}(\lambda) \\ \bar{i}(\lambda) \end{pmatrix}$$

From this equation one can see that a variation of the fundamental primary D in the plane PD ($b_3=0$) affects the colour-matching curve $\bar{p}(\lambda)$ and a variation of D in the plane DT ($b_1=0$) affects the colour-matching curve $\bar{i}(\lambda)$. Therefore out of a variation of one of the fundamental primaries a distortion of that fundamental colour-matching curve results, the vector of which lies in the displacement plane of the changed fundamental primary.

If we displace for instance in Judd's fundamental primary system only the fundamental primary D in the plane PD ($b_3=0$) we get other fundamental colour-matching functions $\bar{p}^*(\lambda)$:

$$\bar{p}^*(\lambda) = \bar{p}(\lambda) - \frac{b_1}{b_2} \bar{d}(\lambda)$$

as shown in Fig. 2. In this case the intersection point of the $p^*(\lambda)$ and the $d(\lambda)$ curve remains.

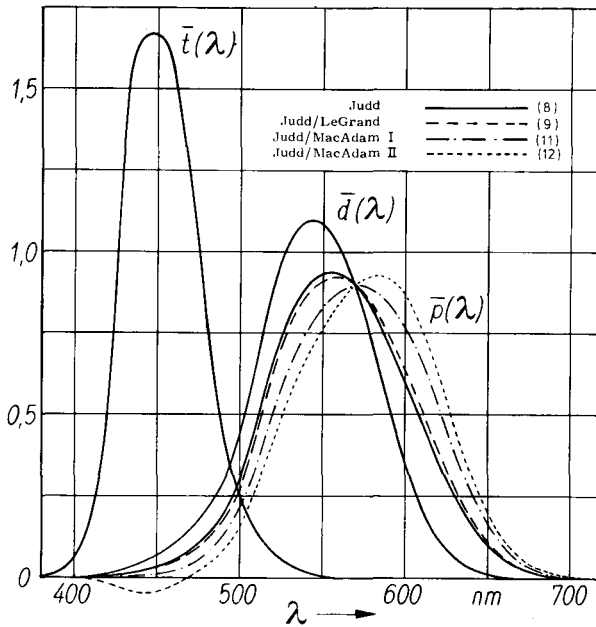


Fig. 2. Fundamental response curves $\bar{p}(\lambda)$, $\bar{d}(\lambda)$, $\bar{i}(\lambda)$ due to various locations of the D -primary in the plane $T=0$. Numbering refers to the coding of Fig. 1.

Changing the illuminant for surface colours it is well known, that the human eye adapts to the new light source that this illuminant will be seen nearly achromatic as well. This chromatic adaptation (which causes an adaptive colour shift) works against the colorimetric shift and neutralizes it in the ideal case. According to von Kries (1902) it can be easily described in the fundamental primary system by a diagonal matrix with the coefficients A , B , C . In the CIE standard colorimetric system the chromatic adaptation matrix U has 9 coefficients and depends on the matrix K , which transforms the tristimulus values from the fundamental primary system into the CIE standard colorimetric system:

If

$$\begin{pmatrix} P \\ D \\ T \end{pmatrix} = K \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

the equation

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_A^* = K^{-1} \begin{pmatrix} A & & \\ & B & \\ & & C \end{pmatrix} K \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_A'$$

or

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_A^* = U \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_A'$$

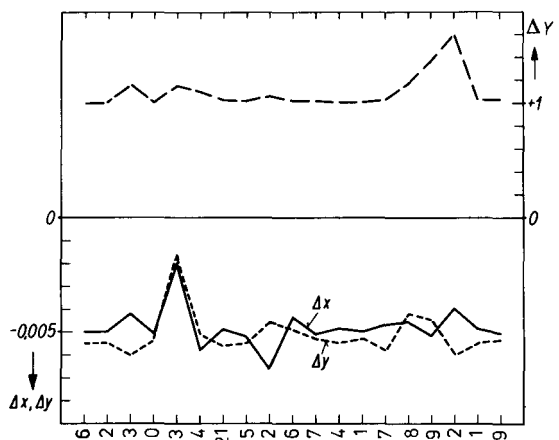


Fig. 4. Difference of Δx , Δy , ΔY resulting from the tristimulus values X_{13} , Y_{13} , Z_{13} of the CIE test colour sample 13 under illuminant A and the tristimulus values $X_{13}+1$, $Y_{13}+1$, $Z_{13}+1$, transformed for chromatic adaptation from illuminant C to illuminant A , using the fundamental systems of the authors. Numbering corresponds to Fig. 1.

Figs 5 and 6 show the colour points of the CIE test colour set for illuminant A and the corresponding colours, which have the tristimulus values of the CIE test colours with an additional 1,0 in XYZ. Chromatic adaptation from illuminant C to illuminant A has been taken into account by transformation matrices from the fundamental primary systems like Fig. 4. One can see that the given colour differences will be shifted by this transformation but the transformed colour differences approximately keep their direction.

To get a better statement for the influence of the fundamental primaries on colour differences with chromatic adaptation from illuminant C to illuminant A we chose a grey colour ($x=0.310$; $y=0.316$; $Y=10$) and a red colour ($x=0.670$; $y=0.300$; $Y=10$). For the two theoretical colours we calculated the

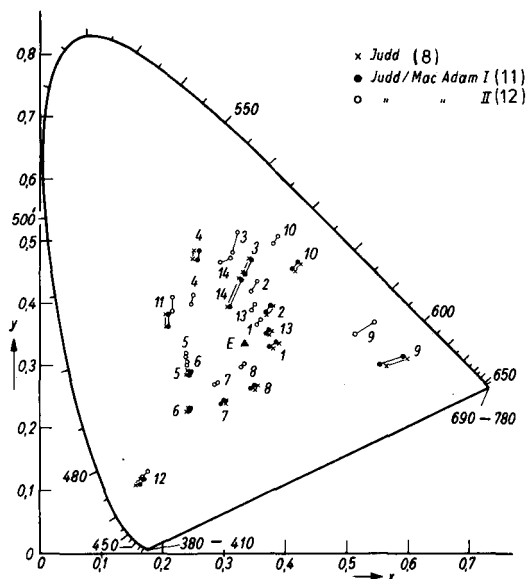


Fig. 5. Calculated colour points of the CIE test colour samples and corresponding colours with an additional 1 for the tristimulus values XYZ under illuminant A and chromatic adaptation from illuminant C to illuminant A , using various fundamental systems. Numbering refers to Fig. 1.

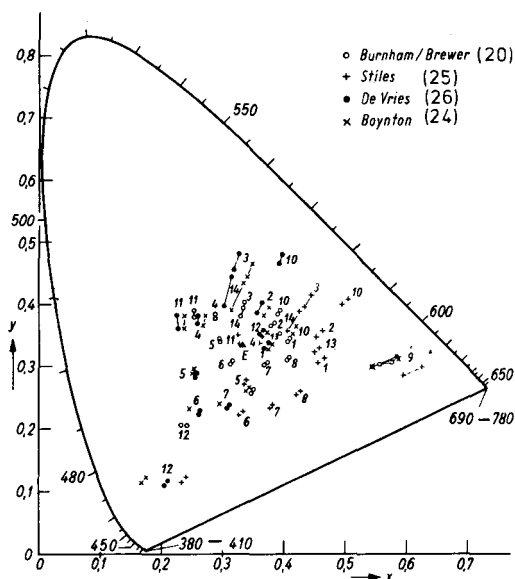


Fig. 6. Calculated colour points of the CIE test colour samples and corresponding colours with an additional 1 for the tristimulus values XYZ under illuminant A and chromatic adaptation from illuminant C to illuminant A , using various fundamental systems. Numbering refers to Fig. 1.

corresponding MacAdam ellipse for a colour difference $\Delta E_{MA}=4$. These ellipses have been transformed by means of the chromatic adaptation matrix of several authors fundamental primary systems. The figures 7 and 8 show that the ellipses resulting from various formulae for the adaptive colour shift have nearly the same shape and size and only their position differs in the chromaticity diagram.

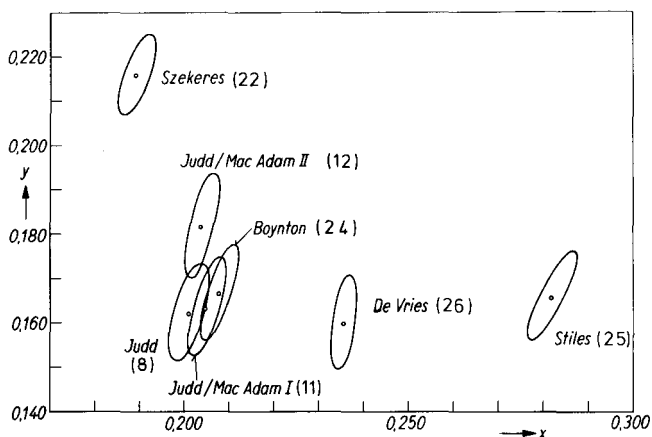


Fig. 7. Theoretical colours, forming an ellipsoid in the colour space with $\Delta E_{MA}=4$ round grey colour $x=0.310$, $y=0.316$, $Y=10$ transformed for chromatic adaptation from illuminant C to illuminant A . Numbering refers to Fig. 1.

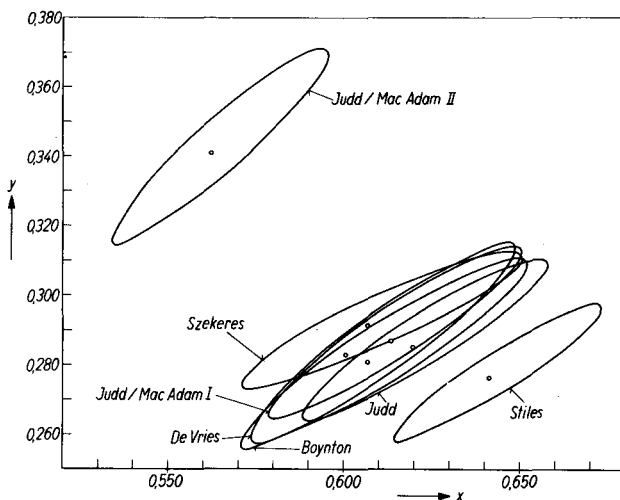


Fig. 8. Theoretical colours, forming an ellipsoid in the colour space with $\Delta E_{MA}=4$ round the red colour $x=0.670$, $y=0.300$, $Y=10$, transformed for chromatic adaptation from illuminant C to illuminant A . Authors as in Fig. 7.

Compared with their previous position the ellipses seem to be distorted a little counterclockwise. Therefore one can assume that different mechanisms of the visual system are responsible for the process of chromatic adaptation and colour discrimination.

From this theoretical investigation one can assume:

1. colour difference formulae which have been generally set up for daylight illumination could be applied to tungsten illumination (illuminant A) in combination with a suitable chromatic adaptation formula.

2. as the chromatic adaptation process of the human eye is assumed to be caused by neurons, the process of colour difference sensation must take place in another phase of the visual system and a line element for colour discrimination cannot be obtained in a linear form from fundamental colour-matching functions.

REFERENCES

- Boynton, R. M. (1969), Theory of colour vision. *J. Opt. Soc. Amer.* 50, 929-944.
 Bouma, P. H. (1942), Colour vision system of trichromats and dichromats. *Physica* 9, 773-784
 Brewer, W. L. (1954), Fundamental response functions and binocular color matching. *J. Opt. Soc. Amer.* 44, 207-212
 Commission Internationale de l'Eclairage Publication no. 13, Method of measuring and specifying colour rendering of light sources.
 Federov, N. T., et al (1953), Über die Grundvalenzkurven für den mittleren Beobachter. *Problemy fiziol. Optiki* 8, 99-111

- Judd, D. B. (1941), Color systems and their interrelation. *Ill. Eng.* 36, 336-372
- Judd, B. D. (1944, 1945), Standard response functions for protanopic and deuteranopic vision. *Bur. of Standard J. Res.* 33, 407-457; *J. Opt. Soc. Amer.* 35, 199-220
- Justova, E. N. (1948), Die Bestimmung der Koordinatenachsen im physiologischen Grundsystem aus Versuchen mit Farbenblinden. *Dokl. Akad. Nauk. SSSR* 63, 383-385
- Justova, E. N. (1950), Die spektrale Empfindlichkeit der Empfänger im Auge. *Dokl. Akad. Nauk. SSSR* 74, 1069-1072.
- Kries, J. von (1902), Theoretische Studien über die Umstimmung des Sehorgans Festschrift der Albrecht-Ludwig Universität Freiburg, p. 144-158.
- Kustarev, A. K. (1965), Über die Grundfarben eines physiologischen Farbsystems. *Sveto-technika* 11, 5-11.
- LeGrand, Y. (1948), Standard response functions for protanopic and deuteranopic vision. *J. Opt. Soc. Amer.* 38, 815-816
- Lobanova, P. A. and Rautian, G. N. (1952), Bestimmung der spektralen Empfindlichkeit der Netzhautrezeptoren aus Versuchen mit Dichromaten, *Dokl. Akad. Nauk. SSR* 146, 1193-1196
- MacAdam, D. L. (1956), Chromatic adaptation. *J. Opt. Soc. Amer.* 46, 500-513
- MacAdam, D. L. (1957), Beat-frequency hypothesis of colour perception. *Visual problems of colour. Proc. Symposium, Teddington*, 578-601
- Nyberg, N. D. and Justova, E. N. (1957), Researches on dichromatic vision and the spectral sensitivity of the receptors of trichromats. "Visual problems of Colour". *Proc. Symposium, Teddington*, 475-486.
- Pitt, F. H. G. (1944), The nature of normal trichromatic and dichromatic vision. *Proc. Roy. Soc.* 132 B, 101-117
- Sperling, H. G. (1960), Case of congenital tritanopia with implications for a trichromatic model of color reception. *J. Opt. Soc. Amer.* 50, 156-163
- Steffen, D., (1955/56), Untersuchungen zur Theorie des Farbensehens. *Z. Biol.* 108, 161-177
- Stiles, W. S. (1946), A modified Helmholtz line-element in brightness-colour-space. *Proc. Roy. Soc.* 58, 41-65
- Szekeres, G. (1948), A new determination of the Young-Helmholtz primaries. *J. Opt. Soc. Amer.* 38, 350-363
- Terstiege, H. (1967), Untersuchungen zum Persistenz- und Koeffizientensatz. *Farbe* 16, 1-120
- Thomson, L. C. and Wright, W. D. (1953), The convergence of the tritanopic confusion loci and the derivation of the fundamental response functions. *J. Opt. Soc. Amer.* 43, 890-894
- De Vries, H. (1946), On the basic sensation curves of the three-color theory. *J. Opt. Soc. Amer.* 36, 121-127
- Walters, H. V. (1942) Some experiments on the trichromatic theory of vision. *Proc. Roy. Soc.* 131, 27-50
- Wright, W. D. (1934) Measurement and analysis of colour adaptation phenomena. *Proc. Roy. Soc. B* 115, 49-87

DISCUSSION

Pointer: You state in your conclusions, in the preprint, that colour difference formulae which have been set up for daylight illumination may be used in tungsten illumination with the addition of a suitable chromatic adaptation formula.

I recently completed a study of the effects of various white light adaptations on the size of the just noticeable colour difference step. I used adaptations with colour temperatures in the range 6500 to 2000 K and found that the change in size of the j.n.d step with changing adaptation was so small as to be considered

negligible. I wonder therefore whether a chromatic adaptation formula is necessary.

Terstiege: If the eye looks to unsaturated lights it will after a time completely adapt to it. Therefore the just noticeable colour differences for illuminants of various colour temperatures must be the same due to the chromatic adaptation.

In contrary to surface colours, in this case a consideration of chromatic adaptation is really unnecessary.

Hunt: This is the first simplifying idea during this conference!

Wyszecki: The x, y -chromaticity diagram is a badly distorted map of chromaticity points from the perceptual point of view. Thus the significance of the spread of the fundamental primaries (particularly D) is very difficult to assess. In fact, I suspect your findings indicate the long expected fact that the spread is not very significant.

MacAdam: Experiments by Brown and myself indicate that chromatic adaptation changes do not influence the shapes or orientations of the CD -ellipses. I think that the ellipse appropriate for the x, y locations of the color (whatever the illuminant of the surround) should be used.

Terstiege: White samples have other chromaticity loci in the CIE diagram with illuminant A than with illuminant C . Without considering chromatic adaptation colour difference formulae will give different results for a white pair with both illuminants. Experience however shows that colour difference evaluation may be done best with a surrounding field of the same colour and therefore the CD -ellipses are smaller in the achromatic point than away from it.

Brookes: I join the point of view of MacAdam from the following direction: if we have very large samples, we have an adaptation to the sample color and should by use of an adaptation formula get another color difference than for the same colors, if they are smaller and in a daylight surround. I find this strange and suppose that the color difference ellipsoid will not be distorted by such adaptation.

Terstiege: Certainly there will be an influence of the surrounding field on the state of chromatic adaptation which will depend on colour and size of this field. For infinite samples therefore the eye will not adapt anymore to the illuminant but to the colour of the surrounding field.

But have you done any experiments of the influence of colour and size of the surrounding field on colour-differences?

Brookes: No.

Hunt: Has anyone done any experiments on this matter?

Simon: Definitely yes, we did. The size of the ellipse changes when very large standards are compared to much smaller samples. Apparently there is no change in the relative axes, nor is there any tilting of the ellipsoids, but this determination is only an observation not confirmed by systematic investigation.

Hunt: Do they change in shape or size?

Wyszecki: In size only.

Stiles: How did you fix the actual amount of chromatic adaptation assumed in your calculations—was it complete in the von Kries sense?

Terstiege: All calculations have been done for complete chromatic adaptation to the illuminant using the linear von Kries formula. A factor for the degree of chromatic adaptation has not been taken into account.

10

GENERAL DISCUSSION I

R. W. G Hunt, Discussion leader

Hunt:

To start the discussion, may I formulate some questions:

1. What is a line-element? How does it differ from a colour difference formula?
2. What are the differences between the empirical approach—championed I think by Dr. MacAdam—and the inductive approach, championed here by Dr. Vos and Dr. Walraven. Can they learn something from one another; is one more useful than the other?

What about the limitations: at the retinal level, in quantum fluctuation, or at more than one stage?*

3. What about the size of the colour differences? Dr. MacAdam likes to deal with a threshold approach, Prof. Wright and others think more in terms of the smallest differences you can set in a field. And those dealing with Munsell samples and with scaling models, work with much bigger differences.

Let us start with the first question.

Wyszecki:

I think Dr. Stiles should be the one to answer that.

Stiles:

I might try to say what I think is the essence of a line element. But I would hesitate to say in what way a colour difference formula is essentially different.

A line element, in my opinion, is an expression which relates the perceptibility of differences between neighbouring colours in a form which should show how these differences become perceptible through different mechanisms of the eye. This is the inductive view, of course. The parameters should, if possible, have some physiological or physical meaning and the element should also be applicable, if possible, to extreme as well as the common conditions met with in practice. It should also be capable of modification in the light of new additions to our ideas of the visual system.

* In this part of the general discussion, this point was not covered at all. However, the issue was picked up again in the General Discussion III (p. 330).

Wyszecki:

The notion of a line element assumes that we accept the idea that we can represent colours by points in a space—usually a 3 dimensional space—and that we can play with it according the rules of mathematics. The geometry need not to be euclidean of course.

In the line element terminology $ds=1$ means one threshold difference. Bigger differences are just the sum of threshold steps. As to smaller differences, I don't think they fit into a line element model without making another assumption, namely that subthreshold steps, e.g. expressed in terms of standard deviations of color matching are a constant fraction of a threshold step.

MacAdam:

The concept of a line element, in my opinion, introduced notions foreign to color and might better be avoided in discussions of color-difference formulas. It prejudices most people against formulas associated with noneuclidian geometry, which formulas are not essentially more difficult and may be more appropriate than the euclidian forms. Positive definite quadratic forms, in general are appropriate for color differences because, at threshold, the effect of a positive increment is equal to the effect of a negative increment of the same kind and amount, from the same starting color. Such behavior can be represented, in general, only by a positive definite quadratic form. It is only a coincidence, and a misleading one, that the same form is also used in Riemannian geometry. Only physiological and colorimetric considerations should govern the formulation of color-difference formulas. Geometry is foreign to them, and has the effect of limiting the options.

Wright:

The term "line element" does not seem to me to have any visual meaning at all and to have no obvious connection with the visual processes responsible for colour discrimination. I also find it objectionable because it encourages the discussion of colour discrimination in purely mathematical terms. For example, in Dr. Stiles' very impressive lecture this morning, I do not think he once referred to the colour sensations themselves, and yet without the generation of redness, greenness, etc. we would have no colour discrimination at all.

Stiles:

Just some 10-20 years before Helmholtz published his line element concept, Riemann had been developing analytical geometry and used this expression for the distance measurement. It had not essentially to do with colour, of course. Helmholtz, himself a great mathematical physicist as well, referred to Riemann in his treatise and took the term from it. It has a drawback, of course, because it tends to make us think of colour differences in terms of a sum of squares or

at least as a quadratic form—which is not experimentally certain. I do not think, however that a more elaborate form is likely to help much.

Wyszecki:

It satisfies our desire to put some mathematical formalism into our world of color.

Wright:

I think this is a disadvantage. It tends to conceal the physiological and perceptual aspects of the problem.

MacAdam:

And yet each maker of systems—Oswald, Munsell etc.—has thought of colour in terms of a mathematical model. It is almost inherent to it. Riemann (1854) himself wrote:

“The position that the objects of sense and the colors are probably the only familiar things whose specifications constitute a multiply extended manifold”. So this association of colors with points in space is as old as Riemannian geometry—and I think it is inevitable. Now unfortunately, to truly simulate the relations of colors, this geometrical model is going to be awfully complicated. It is not as simple as euclidean space. I have often thought that we might better grow out of thinking in terms of this geometrical model. We have to stay close to the data. We want a formula for color difference and may have to stop cluttering up our minds and the literature with line elements. That’s why I don’t speak, or write or draw graphs of the metric coefficients, g_{ik} , any more.”

Stiles:

I disagree on Dr. MacAdam’s last remark. At the moment I don’t see any prospect of getting away from the quadratic summing concept of the line element as a workable theoretical tool.

Wyszecki:

It is just a mathematical formula, I don’t see why we have to talk about it in terms of geometrical concepts.

Stiles:

No, but I am talking about it as an expression with structure whose terms can be associated with different mechanisms.

Wright:

This illustrates my point. All this mathematics has nothing to do with the colour sensations themselves.

McLaren:

I was assuming that the three dimensional space was inevitable. Dr. Wysecki said "usually 3 dimensions" implying that there might be 4. Dr. Stiles said certainly not, he wanted to restrict himself to 3 dimensions. Could either or both of you elaborate a little on this and perhaps make it even more complicated?

Stiles:

I can't recall having said it this way. Without remembering the context, I can only say that I don't subscribe to this statement.

Wysecki:

Dr. Trezona this morning indicated that indeed a fourth, (scotopic), component might come in.

Bartleson:

Is not there a basic semantic problem? What is color? The OSA has defined color as that condition of match which is represented by a tristimulus specification. I think that Dr. Stiles is not limiting his use of the word color to mean simply that definition and I am quite sure that Prof. Wright is thinking about that which we perceive.

Nimeroff:

The problem we have been discussing in this session is a philosophical one in which there is the real world of stimuli and responses on one hand and the abstract mathematical world on the other. In order to use the abstract mathematics to help understand the real world we need to establish the interfacing between these two worlds.

Stiles:

I imagine this points right to Dr. Wright's question.

He complains that we have not talked about sensation. But we don't need to talk about sensation.

Kowaliski:

The question about mathematics has been answered a long time ago by the famous french mathematician and physicist Henri Poincaré (1889). To quote from the introduction of his "Théorie mathématique de la lumière":

"... It cannot be the purpose of mathematical theories to reveal to us the true nature of the physical world; this would be an unreasonable pretention. Their only purpose is to coordinate the physical laws made apparent by experi-

ments, laws which it would be impossible to describe without the help of mathematics . . . ”.

Hunt:

I think Dr. Stiles should explain how he can study colours and dispense with sensations.

Stiles:

You put a man down behind a colorimeter, you guide his hand to three knobs and let him go ahead.

Hunt:

This tells you everything?

Stiles:

Of course not. But you ask him to make certain settings based on the appearance of the colorimeter field. You draw your conclusions from the relations between the stimuli exposed in the fields, and the settings he makes. In expressing these relations it is not necessary to claim one is “measuring a sensation” or in fact to “regard a sensation” as having any particular meaning as a scientific term. Of course, the word “sensation” may be used colloquially to help explain to the observer what you want him to do.

Hunt:

What about magnitude scaling?

Stiles:

Yes, again you ask a description, an external response, and you correlate the responses with the stimuli not with some hypothetical entity “sensation” located in some sense between the two.

Ishak:

There is a dilemma between colour differences and colour difference equations. People like Dr. Stiles and Dr. Nimeroff forget that we have to deal with colour perception and not with mathematical models. There is so much evidence that summing of ΔE 's is not-linear, in particular in chroma-steps. Colour is what we see!

Hunt:

May we now turn to the last question of my list:
What are the most important colour differences, the just noticeable differences or the supra threshold differences like e.g. those in the Munsell system?

Billmeyer:

We are looking to see if there is a fundamental difference between the Munsell and the MacAdam-data. We will, in the near future, make a more thorough investigation comparing these data by studying subsamples of Munsell steps, starting with near-whites.

Friele:

By analyzing small and large colour differences there appears to be a big difference in the scaling principles. I may show that on the basis of Fig. 3 of my paper, yet to be presented. It shows that the dependence of ΔB on B and Y is just the reverse for threshold and for Munsell Δ chroma 2.

Wyszecki:

That is not experimental evidence; instead it is an interpretation of some data in terms of assumed B and Y mechanisms. Moreover, it is impossible to use the Munsell system in this way. First, it has to be established that the differences in this system are perceptually equal.

Nickerson:

Once you go back to the 1943 Newhall-Nickerson-Judd report you will see that that was the intention both of the system, and of the renotation data which is recommended in the OSA 1943 subcommittee report.

There should be no fundamental difference between Munsell and JND scales. Certainly more work is needed to evaluate this, but based upon the way the Munsell system was developed at least no discontinuity was found between small and large color difference steps. The ability to divide the Munsell steps down to their smallest useful size has been one of the useful features of the Munsell-scales.

It seems taken for granted in many discussions at this conference that Munsell scales are derived principally from studies of large-color-differences, those about 2 Munsell chroma steps. Perhaps this is because intervals of this size were the basis for the reports of the OSA subcommittee studies (Newhall *et al*, 1940; 1943) I would like to call attention to the fact that a very considerable amount of small-difference work went into the studies on which the scales of the *Book of Color* (Munsell, 1929) were based, also to the fact that the decision to produce charts for this edition in 2-step chroma intervals, instead of the 1-step intervals of the original Munsell *Atlas*, was made only after the 1924-26 experimental work was completed.

While I had no responsibility for the research carried on in the mid-20s in the Baltimore laboratory, I was there in the business office until late 1926 and in the earlier years occasionally acted as observer in the extensive experiments carried on to develop a revision of the 1912-1915 *Atlas* scales, a revision under-

taken in Baltimore in a laboratory set up for the purpose by A. E. O. Munsell under the guidance of Irwin G. Priest, head of colorimetry at the Bureau of Standards. Priest (1920) had advised that: "A revised edition of the *Atlas and A Color Notation*, based upon the best present-day methods of measurement and specification, would be a most important contribution to the science and art of chromatics generally."

Experiments, beginning in early 1924, were carried out first by A. E. O. Munsell and Miriam E. O'Brien, a Bryn Mawr student of Ferree and Rand recommended by Priest, followed (1925-6) by Louise L. Sloan, a classmate. The names of other experimenters and observers are included in the first of two published histories of the Munsell Color System. (Nickerson, 1940; 1969). Among them were Prentice Reeves, who spent the summers of 1925 and 1926 in Baltimore, Geraldine Walker (Haupt) 1925-7, and I. H. Godlove, 1926-33. Prentice Reeves seems principally responsible for planning the JND studies. Deane Judd, based in Washinton at the Bureau of Standards, spent a number of sessions on this work during Prentice Reeves' stay in Baltimore.

During this time I became well acquainted with the term "just-noticeable-difference" for the early work was done on this basis. More extensive studies were made on value than on hue and chroma scales, but this seemed justified in part (as Louise Sloan and I now recall) because the experiments that *were* made tended to show no great difficulty with hue and chroma. Also, the Bureau report proposals had emphasized that work be done on the value scale.

Experimental results by a "just-noticeable-difference" method and a "value-step" method have been reported in detail for the value scale by Munsell, Sloan and Godlove (1933). They concluded that the results were in "excellent agreement" and that "the precision with which an average value scale can be determined by 14 unpracticed observers, each making a single set of determinations by the method of equal value steps, is equal to that attained in the results of 6 observers making 5 to 6 separate determinations by the just-noticeable-difference method, and the precision of the former method for 3 observers for the average of 4 determinations or less is more than twice as good." A further report on this work by Godlove (1933) concerns more particularly the inclusion of background reflectance in equations relating value and reflectance.

While the bulk of the work was done on value, nevertheless considerable study was made of hue and chroma scales, and for this many small-step samples were prepared. I remember in particular the 100-step hue circuit at mid-value/mid-chroma (used later by Farnsworth as a basis for his "100-hue test"), a 50-step hue circuit at maximum chromas, a 60-step value scale, and a many-stepped chroma scale in several hues—I particularly remember the red scale of 27 (or 33?) steps. Measurements of several of these scales were later made and reported (Granville et al, 1943: Tables II and III). Abstracts for 1926-35 in the Optical Society journal show that several papers relating to this work

were given at meetings of the Society by various authors. Unfortunately only the value scale reports were completed for publication.

Only after the experimental work was completed in 1926 was it finally decided that papers would be painted to illustrate the proposed revision. For the first edition of a new Munsell *Book of Color* papers were painted for only 20 hues, in value steps 2/ through 8/, in even steps of chroma, omitting chroma 2 for 10 of the 20 hues. Although the need for additional hues soon became apparent, no similar need arose for inclusion of the odd-numbered chroma steps. Today papers for 40 hues and many in-between colors are regularly available, e.g., 1 chromas for 20 hues, 3 chromas for *R* to *Y* hues, and several selections of intermediate hues. In fact, it is an important service of the Munsell company to provide papers on request to represent notations falling between those in the stock collection; for tolerance sets these are *very-small-differences*.

I hope that this explanation may help to clarify some of the misconceptions that exist regarding development of the Munsell scales. It should be noted that today the renotation specifications recommended in the final report of the OSA subcommittee on spacing of the Munsell colors (Newhall et al, 1943) provide the standard used by the Munsell Color Company in producing the Munsell papers, now available in glossy as well as matte surface. The aim of the subcommittee was "to make, as nearly as feasible" the loci for hue and chroma "perceptually equi-spaced." If and when sufficiently important improvements on the 1943 specifications can be recommended, no doubt the standard will be adjusted accordingly, for it is the continuing purpose of the Munsell system that it represent a color space describable in terms of equi-spaced scales of hue, value and chroma.

Hunt:

Thank you very much, Dr. Nickerson for this elucidating historical account.

*Judd:**

Although relatively few data were collected I never seriously doubted that for practical purposes, information obtained from large differences, would apply to the spacing of colors showing small differences. For example, in my study of a saturation scale for yellow colors (Judd, 1933).

I showed that change of the step size from one to four JND's made no difference. This does not prove that changing step size by a factor of 40 would be negligible, nor does it mean that change in step size by a factor of 4 is negligible for all kinds of color change. It has only been in the last 10 years when I realized that the Munsell-Sloan-Godlove work on value scales is substantially in error because it does not show crispening, that I have been convinced that a

* Added to the discussion by correspondence.

considerable research investment on the effect of step size is called for.

The conditions used in the JOSA report noted are far from those used in industry, and the one color series studied has no simple relation to Munsell parameters. Nevertheless from this and from even less extensive bits and pieces of studies that are far from conclusive, I have gradually fallen into the view that if we could improve the spacing of the colors in the Munsell book so that no obvious nonuniformities in gross spacing could be seen, then equal fractionation of these equal intervals would yield intervals that likewise appeared to be of equal size, thus justifying use of the Munsell system for size specification of color differences of sizes varying over a wide range, say from 0.4 to 4.0 chroma steps, or a factor of 10. It is high time that we look more carefully into this important question.

MacAdam:

In the consideration of whether small-difference formulas are relevant for evaluation of large color differences, the question whether a difference of $2N$ small units appears twice as great as a difference of N such units is irrelevant: We know that the answer is negative. The proper question is whether a difference of N small units between two colors appears just as great as a difference of N of the same small units between any other two colors, however different they may be from the first pair (N being 10 or greater).

Jaekel:

Coming back to Prof. Hunt's question, I feel there is no simple answer: mechanisms of colour vision generally seem to be investigated by using very small differences, to measure discrimination. Those of us concerned with television reproduction must look at relatively large differences.

Those, who, like myself, are concerned with pass-fail decisions on dyed textiles, are interested in intermediate differences. I suspect strongly that what is equal in one of these domains, may not be equal in another: we are dealing with several sorts of equality, and in the intermediate area our problem is that the majority of colour difference formulae available for us to try out, are based either on much smaller or on much larger differences than those of most practical interest to us.

Reilly:

Is there then a discontinuity? In my opinion it never can be a discontinuity, because adding small color differences is a continuous process.

Billmeyer:

Whether that is true or not is just purpose of our investigation. Presently we are not able to answer this question.

Indow:

1. Threshold is a statistical concept and if we define line-element in terms of threshold, we must deal with the probability of detecting. It is a fundamental question whether or not we shall be able to have the concept of a line-element which is independent of the probability. I would like to mention that a very thorough analysis of that problem has been made, from a mathematical point of view, by Prof. Luce, for example (Luce, 1963)
2. The Munsell system is based on the matching procedure. Hence, as pointed out by Dr. Wyszecki, it has nothing to do with threshold by itself. However, the same intervals in Munsell space should contain the same number of line-elements if counted along the proper geodesics. Otherwise, the term "line-element" loses its meaning.

Hunt:

And that sounds like rounding up the symposium theme. It gets late—and this seems a proper place to stop, indeed.

REFERENCES

- Godlove, I. H. (1933), Neutral value scales. II. A comparison of results and equations describing value scales. *J. Opt. Soc. Amer.* 23, 419-425
- Granville, W. C., Nickerson, D., and Foss, C. E. (1943), Trichromatic specifications for intermediate and special colors of the Munsell system. *J. Opt. Soc. Amer.* 33, 376-385
- Judd, D. B. (1933), Saturation scale for yellow colors. *J. Opt. Soc. Amer.* 23, 35-
- Luce, R. D. (1963), Detection and recognition. In: Luce, R. D., Bush, R. R., and Galanter, E. (eds.), *Handbook of mathematical psychology*. Wiley, New York, Chapter 3.
- Munsell book of color, 1929 (in 1970 available bound, loose-leaf or file), Munsell Color Co., 2441 North Calvert Street, Baltimore, Md. 21218, USA.
- Munsell, A. E. O., Sloan, L. L., and Godlove, I. H. (1933), Neutral value scales. I. Munsell neutral value scale. *J. Opt. Soc. Amer.* 23, 394-411
- Newhall, S. M. (1940), Preliminary report of the OSA subcommittee on spacing of the Munsell colors. *J. Opt. Soc. Amer.* 30, 617-645
- Newhall, S. M., Nickerson, D., and Judd, D. B. (1943), Final report of the O.S.A. subcommittee on the spacing of the Munsell colors. *J. Opt. Soc. Amer.* 33, 385-418
- Nickerson, D. (1940), History of the Munsell color system and its scientific application. *J. Opt. Soc. Amer.* 30, 575-586
- Nickerson, D. (1969), History of the Munsell color system. *Color engineering* 7, 5: 42-51
- Poincaré, H. (1889), *Théorie mathématique de la lumière*. Carré, Paris
- Priest, I. G., Gibson, K. S., and McNicholas, H. J. (1920), An examination of the Munsell color system. *Bur. Standards Technologic Paper no. 167*. 33 pp.
- Riemann, B. (1854) Quoted by D. L. MacAdam in "Sources of Color Science", p. 61. MIT. Press, 1970.

11 MULTIDIMENSIONAL MAPPING OF SIXTY MUNSELL COLORS BY NONMETRIC PROCEDURE

TAROW INDOW AND KIMIKO OHSUMI

*Keio University,
Tokyo, Japan*

The basic postulates and the spacing of the Munsell space will be examined with the aid of multidimensional scaling (MDS). Through MDS we can construct a configuration of n points in a m dimensional space Ω_m in such a way that the interpoint distances d_{jk} between any pair of points P_j and P_k represent the subjectively scaled color differences d_{jk} between the two colors j and k ($j, k = 1, 2 \dots, n$). We have two different procedures of MDS: metric and non-metric (Torgerson, 1958; Shepard, 1962a, b; Kruskal, 1964a, b). Indow and his collaborators first applied metric MDS and then nonmetric MDS to sets of n Munsell colors where n changed from 9 to 24 and the colors varied either in Value, V , and Chroma, C , or in Hue, H , and C , or in all the attributes. (Indow and Shiose, 1956, 1958; Indow and Uchizono, 1960; Indow and Kanazawa, 1960; Indow, 1963, and in preparation). The present application has the following features.

1. In the previous studies, perceived color differences d_{jk} of all possible pairs of n colors were used and the matrix D of all n^2 perceived differences was complete. Hence, when n colors scatter widely in the Munsell space, the subject (S) had to assess extremely large color differences in addition to color differences which are intuitively tangible and easily differentiated in degree. For example red (5R, 10/5) and green (5G, 8/5) are simply "extremely different" and their difference exceeds the capacity of the S to scale the degree of difference among them. On the other hand, when n colors within a local region of the Munsell space were used, it was discovered that the S s seem to change their mode of assessing color differences and anomalous results were obtained (Indow and Matsushima, 1969). The overall differences between two colors were overwhelmingly determined by their difference in V and the difference in H lost its weight. In other words, the spacing of colors for the S is not independent from the range of color differences to be presented in a session. Hence, 60 colors scatter-

ing all over the Munsell space were used in the present study but judgment of color difference was made only with such pairs of colors that are intuitively assessable. Of 1770 possible pairs of 60 colors, the empirical values of d_{jk} were obtained with 419 pairs only and the matrix D , 60×60 , was incomplete.

2. In order to scale the perceived color difference δ as a latent variable and to represent δ as a manifest variable d , the S has to compare many pairs of colors and make some kind of judgment. Ultimately, the judgment is made in terms of either *difference* or *ratio* between color differences. That is true even when matching is made between color differences, δ_{ij} and δ_{kl} , because the matching does not lead to the scale value d_{ij} unless we have in advance a scale value d_{kl} for the other color difference δ_{kl} and d_{kl} has to be scaled ultimately on the basis of either difference or ratio judgment. In scaling of a sensory attribute of a stimulus, e.g., brightness or loudness, it was shown by Stevens (e.g., 1961) over a dozen cases that judgments in terms of difference or ratio between two stimuli do not lead to one and the same scale. That will be true with scaling of color difference and the judgment in terms of difference or ratio between two pairs of colors will yield two different scales which are not linearly related to each other. In the previous applications of *MDS*, metric as well as nonmetric, when the color differences d_{jk} scaled through ratio judgment were plotted against $\hat{d}_{jk}$, the interpoint distance of the obtained configuration, then, a curvilinearity, concave upward, was always observed especially in the region of smaller $\hat{d}_{jk}$. With the purpose of eliminating the curvilinearity, in the present study will be used d_{jk} that are scaled on the basis of difference judgment in the sense to be stated below.

EXPERIMENT

Color Stimuli Each stimulus was 0.8×0.8 cm in size and pasted on a gray card (5.2/) of 2.2×2.2 cm, both being mat on the surface. The stimuli were of two groups. One group consisted of 60 colors as given in the left side figures in Fig. 2 and covered the major portion of the Munsell space. Three gray stimuli were included in this group. The other group consisted of 26 grays with values varying from 3.0 to 8.0 with steps of 0.4 V . The latter was used for scaling color differences among the 60 colors.

Procedure. From the 60 colors, a standard color j was selected and colors k of which differences from the color j were to be scaled. The colors were 5~8 in number and S was asked first to place them in a row according to the order of perceived differences from the color j . Between the S and the row, there were a standard gray g_0 and the 26 grays, and g_0 was one of the four, 2.0, 4.0, 6.0, and 8.0 in V . The S was asked to place the standard color j and the standard

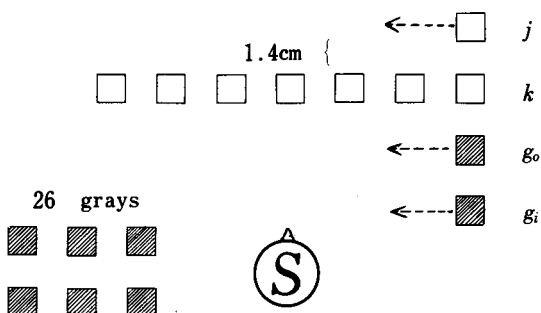


Fig. 1. Presentation of stimuli

gray g_0 in line with the color k at one end of the row; sometimes the most similar color to j and sometimes the most dissimilar color to j . Then, the S selected a gray g_i , one by one, from the 26 grays and placed it on his side of g_i until the overall difference between the colors j and k and the lightness difference between the grays, g_0 and g_i looked the same in magnitude. The S was allowed to select a gray g_i as many times as he wished. In short, matching of the color difference with the lightness difference was made between two pairs of stimuli, the four being placed in a row toward the S (Fig. 1). Two stimulus cards in a pair were placed side by side so that two colors or grays were separated by the gray card (5.2/) of 1.4. cm width. When the above described procedure was completed with the color at the end, then the standard color j and the standard gray g_0 were placed in line with the next color in the row and the whole procedure was repeated with this color. In this way, the color difference with the standard j and each color in the row was matched with the lightness difference. Note that the Munsell V is one of the most well established scales and the extensive data underlying the scale are all based upon difference judgment (Indow & Kawai, 1960). Of each pair of color stimuli, j and k , the matching was repeated twice; once j as the standard with one of the four grays as g_0 and once k as the standard with a different gray as g_0 . The obtained two lightness differences in terms of V were averaged to give d_{jk} for the S . For darker and lighter standard grays g_0 , lighter and darker grays were used respectively as the other member g_i to make up the lightness differences. The total number of matchings for a S was 419×2 and about 10 sessions were needed for a S to complete the whole experiment.

All the stimuli were presented on a large sheet of gray paper (5.2/) and the standard color and the other colors, 5~8 in number, as well as a pair of grays were observed directly in the beam of the standard illumination C of 260 lux. The stimulus cards were illuminated perpendicularly and observed by the S at the angle of 45° from a distance of about 40 cm.

Subjects. Five Ss, all normal in color vision, were used. Two were very experienced, two had some and one had no previous experience in this type of observation of color difference.

Remarks. In the instructions it was emphasized that the overall impression of difference between colors was to be transformed to the lightness difference. It became apparent in the preliminary test that the Ss tend to (separate) the lightness difference between two colors and equate the difference between two grays to this difference in lightness aspect rather than to the overall color difference. Hence, it was urged from the beginning to refrain from this tendency. Still it is doubtful that all the Ss were successful in strictly following the instructions because to match the difference of two colors in the attribute of lightness with the lightness difference between two grays were especially tempting when the corresponding members of the two pairs, colors and grays, were respectively of about the same absolute levels in lightness. Such an annoying pair of pairs was avoided as far as possible.

DATA PROCESSING

As stated above, four matchings at different absolute levels of lightness for a pair of colors j and k were averaged to give $d_{jk,i}$ where i denotes the S. Then, the average over the 5 Ss was defined as d_{jk} . Five separate plots, $d_{jk,i}$ against d_{jk} were made to see whether there was any eccentric S whose judgments are radically different from those of the others. In each plot, the scatter was approximately of the size that could be expected from the scatter in the plot, $d_{jk,i}$ against $d_{kj,i}$. Consequently, d_{jk} defined on the basis of the 5 Ss were put into analysis.

Denote by Ω_m a m -dimensional space in which a configuration of points P_j or P_k is constructed through MDS from the given data D . In metric MDS, Ω_m is Euclidean by definition and the interpoint distance $\hat{d}_{jk}$ is made as close to the given d_{jk} as possible. In nonmetric MDS, if necessary, we can conceive Ω_m to be some kind of non-Euclidean space. However, because it has been shown that conceiving Ω_m to be Euclidean is appropriate in case of colors (Hyman & Well, 1967, 1968; Kruskal, 1964a), and that putting $m=3$ is adequate with the set of colors as used in the present experiment (Indow & Kanazawa, 1960; Indow, 1963), such a configuration was constructed through nonmetric MDS in Euclidean Ω_3 that fitted the data best in the following sense. Denote by $\hat{d}$ a monotonic function of $\hat{d}$: $\hat{d}_{jk}=f(\hat{d}_{jk})$, then such a set of configuration in Ω_3 and monotonic f is looked for through iterations by nonmetric MDS where $\hat{d}_{jk}$ reproduce d_{jk} as closely as possible. We have a number of different procedures, each in the form of a computer program (Guttman, 1968; Lingoes, 1968;

Kruskal, 1964a, b; Shepard, 1962a, b). In the present study, the Kruskal program was used.

For the sake of saving computer time and of preventing the iterations from stopping at a local minimum, the stimulus configuration (left side figures in Fig. 2) was used as an initial configuration, and an understandable configuration was obtained in the Euclidean Ω_3 (the right side figures in Fig. 2).

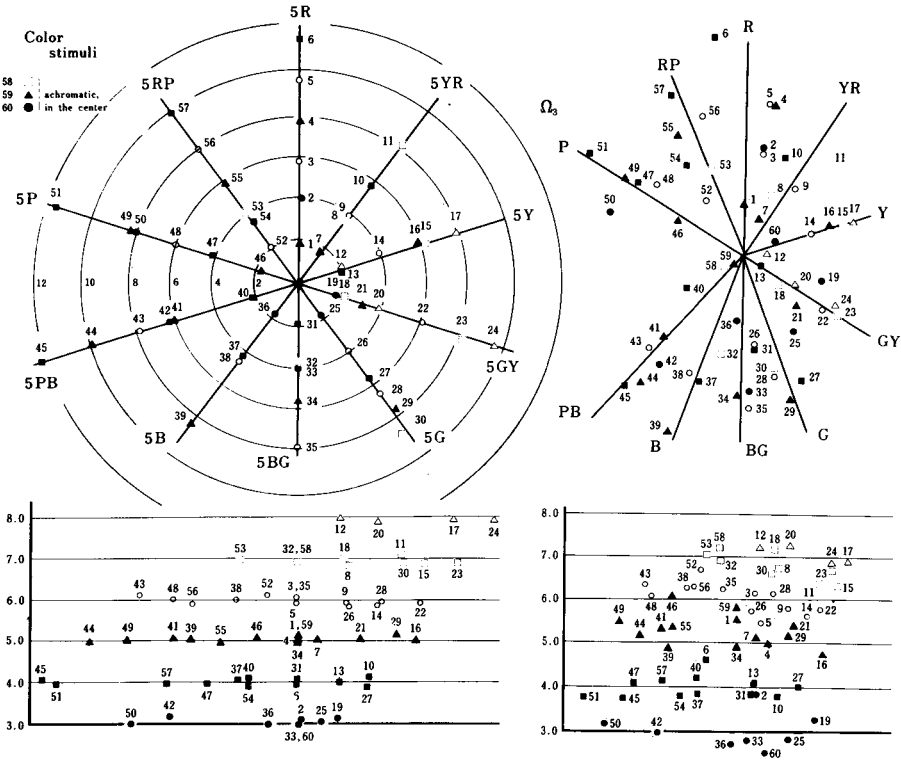


Fig. 2. Color stimuli in the Munsell space (left), and the configuration obtained by multi-dimensional scaling (right). Specifications are in terms of Munsell Renotation.

RESULTS

The goodness of fit of the model to the data can be evaluated by an index which is based upon the scatter of points around the monotonic curve $\hat{d}$ in the plot of d against $\hat{d}$:

$$\text{Stress} = \left[\frac{\sum_{j < k} (d_{jk} - \hat{d}_{jk})^2}{\sum_{j < k} d_{jk}^2} \right]^2 \quad (1)$$

The scatter diagram is shown in Fig. 3 and the stress was 0.069, which implies the goodness of fit of medium degree. As is clear from Fig. 3, the data d and the results $\hat{d}$ are almost linearly related. Major findings will be enumerated below.

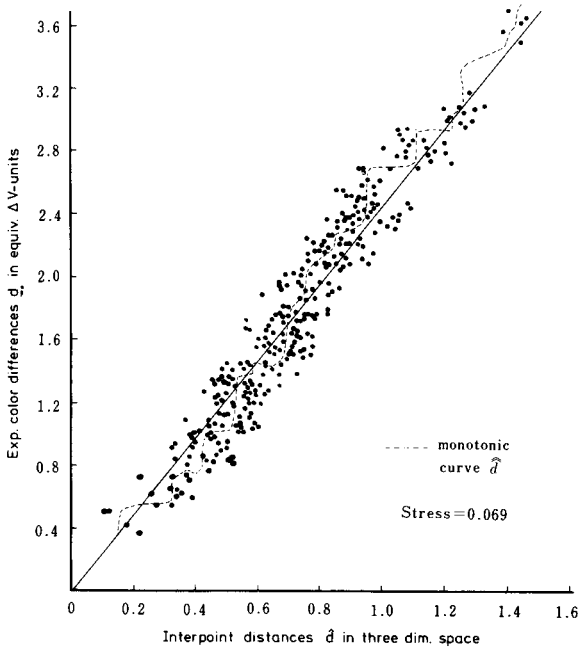


Fig. 3. Relation between experimental data d and results $\hat{d}$, obtained by multidimensional scaling.

$\hat{d}$: monotonic relationship on which the calculation of stress is based

1. In principle, the Munsell space can be regarded as Euclidean and the conclusion is in accordance with the results of other studies (Hyman & Well, 1967, 1968; Kruskal, 1964a), and the axis corresponding to the Munsell value is orthogonal to the plane of constant lightness, not by definition but as a consequence of the analysis.

2. The colors of the same H , even if differing in V , appeared to lie more or less along a line segment approximately in the order of their values of C , and the line segments converged into a single point at the center which represents achromatic color.

3. The six different levels in V from 3.0 to 8.0 were identified in the results though the separation between the two levels, 7.0 and 8.0, was not quite clear.

4. Of the 3 gray stimuli, 58 to 60, two were located in the vicinity of the central point (58, 59) and two were located at the corresponding lightness levels (58, 60).

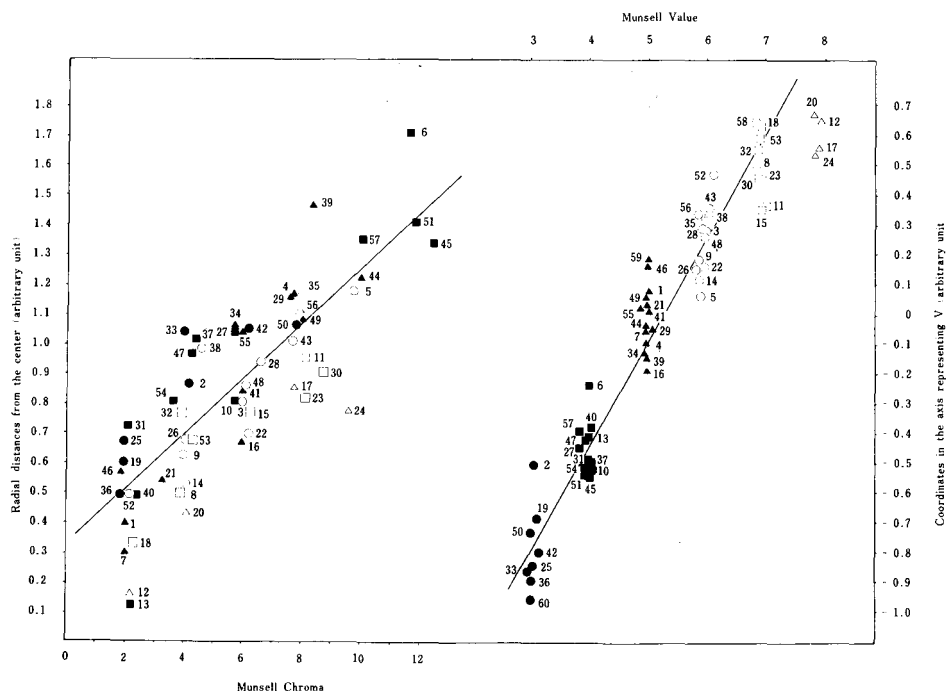


Fig. 4. Correspondences of obtained coordinates with Munsell V and also C respectively.

5. Two plottings are shown in Fig. 4: coordinates of P_j in the axis representing Equivalent Value versus Munsell Value of color j and radial distance of P_j from the center (Equivalent Chroma versus Munsell Chroma) of color j . Scatter in the vertical direction is larger in C than in V and it is for this reason that the 6 equilightness planes with a constant separation from one to the next were introduced as the lower figure of Fig. 2 whereas it was rather difficult to introduce into the upper figure of Fig. 2 equichroma curves as concentric circles with a constant interval. However, from the slopes of two straight lines in Fig. 4, it can be concluded that, generally speaking, 1.0 in Munsell V is subjectively equal in size to 2.8 in Munsell C . The ratio is of about the same order of magnitude as that found in other studies (Indow, in preparation).

6. Pairs of radius vectors, representing two complementary hues are exactly in the opposite direction for $(R:BG)$, $(RP:G)$, and $(P:YG)$, and almost so for $(YR:B)$. It is not true, however, for $(PB:Y)$, which is another way of saying that BG , B and PB are too close together whereas P and PB are anomalously separated. The anomaly of the same sort in the same region, from B to P through PB , has been repeatedly observed in the previous studies with Japanese Ss (Indow & Uchizono, 1960; Indow & Kanazawa, 1960; Indow, 1963). However, it seems not plausible to ascribe the anomaly to their peculiarity because

the same tendency is noticeable in the results of Ramsey (1968) and even of Ekman's data (1954), analysed by Shepard (1962b), in which monochromatic lights were used instead of object colors. Something seems to be wrong with the Munsell spacing of H in this region.

CONCLUSION

All the findings listed above are in complete agreement with those obtained in the previous applications of MDS in the Munsell space (Indow, 1963, and in preparation). Presumably the most important findings are the anomaly of hue spacing in the region from B to P , and the relative sizes of two Munsell units in V and C . Applying MDS seems especially useful to acquire information about the macroscopic properties of the Munsell colorspace. On the other hand, we shall not be able to go into fine adjustment in spacing of each color through this global method of observation on a limited number of colors being sporadic in the color space.

ACKNOWLEDGMENT

The authors wish to express their gratitude to Dr. J. B. Kruskal, Bell Laboratories, whose program has been used in this study and also to Dr. Kawakami and Mr. Hirai, Japan Color Research Institute, who provided the color papers. The study was supported by Toshiba Research and Development Center and Electrical Communication Laboratory, NTT, Japan.

REFERENCES

- Ekman, G. (1954), Dimensions of color vision. *J. Psychol.* 38, 467-474.
- Guttman, L. (1968), A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika* 33, 469-506.
- Hyman, R., and Well, A. (1967), Judgments of similarity and spatial models. *Perception & Psychophysics* 2, 233-248.
- Hyman, R., and Well, A. (1968), Perceptual separability and spatial models. *Perception & Psychophysics* 3, 161-165.
- Indow, T. (1963), Two kinds of multidimensional scaling methods as tools for investigating color space from the macroscopic point of view. *Acta Chromatica* 1, 60-71.
- Indow, T., Applications of multidimensional scaling in perception (in preparation for Handbook of Perception being edited by Carterette, E. C. and Friedman, M. P.).
- Indow, T., and Kanazawa, K. (1960), Multidimensional mapping of Munsell colors varying in hue, chroma, and value. *J. Exp. Psychol.* 59, 330-336.
- Indow, T., and Kawai, T. (1960), A construction of uniform lightness scales on various backgrounds in unified terms. *Jap. Psychol. Res.* 2, 1-12.
- Indow, T., and Matsushima, L. (1969), Local multidimensional mapping of Munsell color space. *Acta Chromatica* 2, 16-24.
- Indow, T., and Shiose, T. (1956), An application of the method of multidimensional scaling to perception of similarity of difference in color. *Jap. Psychol. Res.* 3, 45-64.
- Indow, T., Shiose, T. (1958), A note on an application of the method of multidimensional

- scaling to perception of similarity or difference in colors. *Jap. Psychol. Res.* 5, 21.
- Indow, T., and Uchizono, T. (1960), Multidimensional mapping of Munsell colors varying in hue and chroma. *J. Exp. Psychol.* 59, 321-329.
- Kruskal, J. B. (1964a), Multidimensional scaling by optimizing goodness of fit to a non-metric hypothesis. *Psychometrika* 29, 1-27.
- Kruskal, J. B. (1964b), Nonmetric multidimensional scaling: A numerical method. *Psychometrika* 29, 115-129.
- Lingoes, J. C. (1968), The multivariate analysis of qualitative data. *Multivariate Behavioral Res.* 3, 61-94.
- Ramsay, J. O. (1968), Economical method of analyzing perceived color differences. *J. Opt. Soc. Amer.* 58, 19-22.
- Shepard, R. N. (1962a), The analysis of proximities: Multidimensional scaling with an unknown distance function I. *Psychometrika* 27, 125-139.
- Shepard, R. N. (1962b), The analysis of proximities: Multidimensional scaling with an unknown distance function II. *Psychometrika* 27, 219-246.
- Stevens, S. S. (1961), The psychophysics of sensory function. In *Sensory communication*. (ed. by W. A. Rosenblith). Cambridge Mass., M.I.T. press, 1-33
- Torgerson, W. S. (1958), *Theory and methods of scaling*. New York, Wiley

DISCUSSION

Bartleson: Have you compared the supra-threshold relations determined in your experiments with the perceptual limen for Munsell spacing reported some years ago by Newhall and Bellamy?

Indow: No, I have not done it yet. It was my opinion that the results obtained by this type of approach do not have so much information to predict the threshold. However, I have realized recently that it would be worthwhile trying such comparison as suggested by you.

Jaekel: Fig. 2 (bottom half) shows MDS compression of Munsell Value in the 7-8 region relative to the 3, 4, 5 region: you mention the unclear separation of levels 7 and 8 (on p. 7 of preprint). Therefore, on the right hand side of Fig. 4 would it not be better to draw a convex-upward curve through the mid-points of each vertical "ladder" at one Munsell value, to indicate that subjectively you find 6, 7, 8 to be perceptually less distinct than the lower Munsell Values and that the Value scaling is not perceptually uniform according to your results?

Indow: The reason that I did not try to fit a curve, instead of a straight line, is that the Munsell Value was used as the reference in this experiment to quantify all the color differences, and hence no curvilinearity is allowed for the plotting on the right side in Fig. 4.

Wyszecki: The particular selection of pairs of colors to be presented to the observer must be crucial to the outcome of the MDS. I suspect that the best results are obtained when the pairs presented have about the same perceptual size. If the pairs involved have small as well as large differences, the results may be difficult to reconcile with visual experience.

Indow: Responses of subjects certainly depend upon the range of color differences to be presented. In our 1969 paper, I referred to, Matsushima and I

applied MDS to colors in a restricted region of the Munsell color solid and the result was very different from those in the other studies. In the present study, though pairs of colors presented were limited up to a certain size, the colors covered the whole Munsell color space. And all possible sizes of color difference, say from 0.5~20 in $U^*V^*W^*$, were presented to the subjects.

Ganz: Concerning obtained configuration of hue distribution (Fig. 2, top, right side) it seems surprising that the angle (and therefore colour differences) between PB and P is larger than between PB and G .

Indow: It is surprising to me too.

Hard: Could the result of the experiment be affected by the choice of the reference-scale-valuescale. Would it have been different if the reference-scale was an arbitrary, continuous scale?

Indow: No, I do not think so. We took the Munsell Value scale as the reference in the present study, because it is one of the most well established *interval scale*. In our previous studies, there was no reference and a color pair was compared with another color pair. Nevertheless, we obtained the same result.

Hunt: On the correlation of Munsell and $U^*V^*W^*$ spacing: Munsell renotations were converted into $U^*V^*W^*$ specifications in planes of C^* plotted against W^* , where $C^* = \sqrt{U^{*2} + V^{*2}}$ and examined for uniformity of spacing. It was found that if the $U^*V^*W^*$ space is modified as follows:

$$\begin{aligned}U^* &= 13 (W^* + 12) (u - u_0) \\V^* &= 26 (W^* + 12) (v - v_0) \\W^* &\text{ unchanged}\end{aligned}$$

then the uniformity of spacing as judged by Munsell renotation was improved, and was probably good enough for many applications.

Nickerson: I believe that the results you obtained of 1 V step = 2.8 Chroma steps are based on a sample separation of 1.4 cm between the samples as they were observed. Is it not true that the relation can be expected to differ greatly depending on the degree of separation? (In this connection I refer to an early paper by Bellamy and Newhall (1942) and to one by Nickerson and Newhall (1943) that indicates how great the variation can be.)

Indow: Yes, the relative sizes between V and C would have been different if we had placed two colors directly side by side. Through our experience we adopted 1.4 cm separation as a representative condition for not juxtaposed color pairs.

REFERENCES

- Bellamy, B. R., and Newhall, S. M. (1942), Attributive limens in selected regions of the Munsell color solid. *J. Opt. Soc. Amer.* 32, 465-473
 Nickerson, D., and Newhall, S. M. (1943), A psychological color solid. *J. Opt. Soc. Amer.* 33, 419-422

12 DESCRIPTION OF COLOUR ATTRIBUTES AND COLOUR DIFFERENCES

K. RICHTER

*Bundesanstalt für Materialprüfung,
1 Berlin 45, U. d. Eichen 87*

INTRODUCTION

In our earlier developed opponent colour concept (Richter, 1969a, b) the achromatic and chromatic functions $X_i(\lambda)$ ($i=1, 2, 3$) are linear transformations of the *CIE* spectral tristimulus values $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $z(\lambda)$. The chromatic functions $X_i(\lambda)$ have zero points at the wavelengths of the four spectral unique colours and are similarly defined as the chromatic response curves of Hurvich and Jameson (1955). The main difference with this theory is given by the fact, that the chromatic coordinates $X_i(U)$ and the purity coordinates

$$p_i(U) = \frac{\int X_i(\lambda) u(\lambda) d\lambda}{\int \bar{y}(\lambda) u(\lambda) d\lambda} = \frac{X_i(U)}{Y(U)} \quad (i=2, 3) \quad (1)$$

become different from zero for an achromatic surround U of the relative spectral energy distribution $u(\lambda)$. With this assumption we then succeeded in describing the perceived hue as function of purity and luminance (Richter, 1970).

OPTIMAL COLOUR SOLID

We define the components of the chromatic moment

$$M_i(F) = [p_i(F) - p_i(U)] Y(F) = X_i(F) - p_i(U) Y(F) \quad (2)$$

The chromatic moment gets zero, if the colours F and U in the central and surround field have equal purities.

So, on the one hand the chromatic coordinates $X_i(F)$, $X_i(U)$ of achromatic colours F and U are different from zero, on the other hand the chromatic moments M_i are equal to zero. Chromatic coordinates and moments may correspond to two different levels in our visual system. The first represents an

unchangeable "physical" level, the second level considers an influence of the surround U on the central field F .

We get the transformation coefficients between $M_i(F)$ and $Y(F)$, $X(F)$, $Z(F)$, if we use the chromaticity point of CIE illuminant C ($x=0.3101$, $y=0.3162$; $p_2(U)=-1.7907$, $p_3(U)=0.1425$), and the transformation coefficients between $X_i(F)$ and $Y(F)$, $X(F)$ and $Z(F)$ used in earlier papers (Richter, 1969 a, b).

$$\begin{aligned} M_2 &= -0.4139 X(F) + 3.2478 Y(F) - 2.4046 Z(F) \\ M_3 &= 2.9797 X(F) - 2.5237 Y(F) - 0.0960 Z(F) \end{aligned} \quad (3)$$

Our radial chromatic moment is similarly defined as the well known Adams or Luther-Nyberg moments. We use the scaling factors $n_2=1$, $n_3=2.8$ and the radial chromatic purity $p_r(F, U)$.

$$\begin{aligned} M_r(F) &= \sqrt{n_2^2 M_2^2(F) + n_3^2 M_3^2(F)} \\ &= \sqrt{n_2^2 (p_2(F) - p_2(U))^2 + n_3^2 (p_3(F) - p_3(U))^2} Y(F) = p_r(F, U) Y(F) \end{aligned} \quad (4)$$

If we vary the luminance factor of a central colour F in an achromatic surround U , we can determine three perceived outstanding central colours by visual experiments. We combine their terms and their determination principles in Table I.

Table I. Terms and determination principles of C_0 -, G_0 - and L_0 -colours

Central field Colour	colour attribute	determination principle
C_0 -colour	zero chromaticness	between neither chromatic nor achromatic Evans <i>et al</i> (1967): at the chromatic threshold
G_0 -colour	zero blackness	between neither containing black (gray) nor not containing black Evans <i>et al</i> (1967): equal grayed as the surround
L_0 -colour	equal lightness	equal light (bright) as the surround Kowalski (1969): equivalent luminance Hård (1969): minimum colour contrast compared to the surround

We compute the chromatic moments and luminance factors of optimal colours for $\lambda_d=700$ nm and CIE illuminant C by Eqs (3) and (4). The optimal colours compared to the C_0 -, G_0 - and L_0 - colours, which are described by different equations on the following pages, are also shown in Fig. 1.

G_0 -COLOURS

The G_0 -colours for achromatic and chromatic surrounds were determined by Evans and Swenholt (1967, 1968, 1969). We can describe approximately the experiments by the following equation

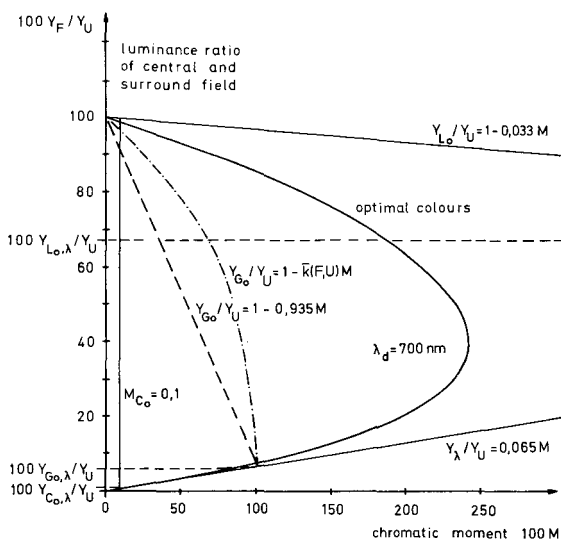


Fig. 1. Chromatic moments and luminance factors of optimal colours, C_0 -, G_0 - and L_0 -colours ($\lambda_d = 700$ nm, $p_r(\lambda_d, U) = 15.4$)

$$\frac{Y_U}{Y_{G_0}(F)} - 1 = |p_r(F, U) - p_s| \cdot \frac{|p_r(F, W) - p_s|}{1 + |p_r(F, W) - p_s|} \quad (5)$$

In this equation the letters F , U and W relate to the central colour F , to the surrounding colour U and to the central colour W , which is perceived as achromatic by chromatic surrounds U . In case of achromatic surrounds (complete adaptation) W is identical to U . The value of the purity threshold $p_s = 0.1$ is normally small compared to radial purity p_r .

G_0 -COLOURS FOR ACHROMATIC SURROUNDS

We neglect p_s in Eq. (5) and get for an achromatic surround ($U = W$) of CIE-illuminant C

$$\frac{Y_U}{Y_{G_0}(F)} - 1 = p_r(F, U) \frac{p_r(F, U)}{1 + p_r(F, U)} \quad (6)$$

Using for short

$$\bar{k}(F, U) = \frac{p_r(F, U)}{1 + p_r(F, U)} \quad (7)$$

Eq. (5) changes in

$$\frac{Y_U}{Y_{G_0}(F)} - 1 = p_r(F, U) \bar{k}(F, U) \quad (8)$$

The following equation corresponding to Eq. (4)

$$M_{G_0}(F) = p_r(F, U) Y_{G_0}(F) \quad (9)$$

leads to

$$\frac{Y_{G_0}(F)}{Y_U} = 1 - \frac{\bar{k}(F, U)}{Y_U} M_{G_0}(F) \quad (10)$$

If the colour F has spectral purity ($p_c = 1$), we must use the letter λ instead of F from Eq. (6) to (10). Please take care that we use the letter λ not only for spectral colours ($\lambda = \lambda_d$) but also for purple colours ($\lambda = \lambda_c$). The luminance factors of G_0 -colours are computed by Eq. (8). The good agreement between the experimental results of Evans and Swenholt (x) and our theory is shown in Fig. 2.

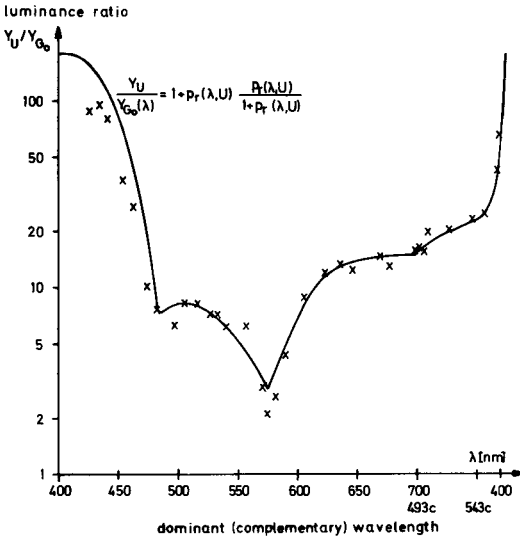


Fig. 2. Luminance ratio Y_U/Y_{G_0} of G_0 -colours for $p_c = 1$, experimental (x) by Evans and Swenholt (1967), computed by Eq. (6).

The change of $\bar{k}(\lambda, U)$ as function of λ is small because $3 < p_r(\lambda, U) < 200$ for all λ 's. That means, the experimental values of Evans and Swenholt (1967) correspond in good approximation to our chromatic purity $p_r(\lambda, U)$.

When we use the spectral chromatic purity $p_r(\lambda, U) = 15.4$ for $\lambda_d = 700$ nm in equations (8) and (9), we get $\bar{k}(\lambda, U) = 0.94$ and $100 Y_{G_0}(\lambda)/Y_U = 6.6$ (compare Fig. 1).

For achromatic central colours ($F = U$), Eq. (6) leads to $100 Y_{G_0}/Y_U = 100$; in agreement with experiments (Fig. 1).

We can use the approximation $\bar{k}(F, U) = \bar{k}(\lambda, U)$ for central colours of any other purity. It follows

$$\frac{Y_U}{Y_{G_0}(F)} - 1 = p_r(F, U) \bar{k}(\lambda, U) \quad (8a)$$

$$\frac{Y_{G_0}(F)}{Y_U} = 1 - \frac{\bar{k}(\lambda, U)}{Y_U} M_{G_0}(F) \quad (10a)$$

The G_0 -colours are located on a straight line (— —) or on a curve (— . —), if we calculate these colours by Eq. (10a) or (10) respectively (see Fig. 1). The differences of both formulations in a uniform colour space are much smaller than in Fig. 1. We must take into consideration the cubic luminance factor scaling for the perceived colour differences.

The G_0 -colours of equal luminance factor are calculated with Eq. (8) and our earlier published transformation coefficients between the purity coordinates (p_2, p_3) and the chromaticity coordinates (x, y) (see Richter, 1969b). Fig. 3 shows that G_0 -colour of equal luminance factor are located on ellipses.

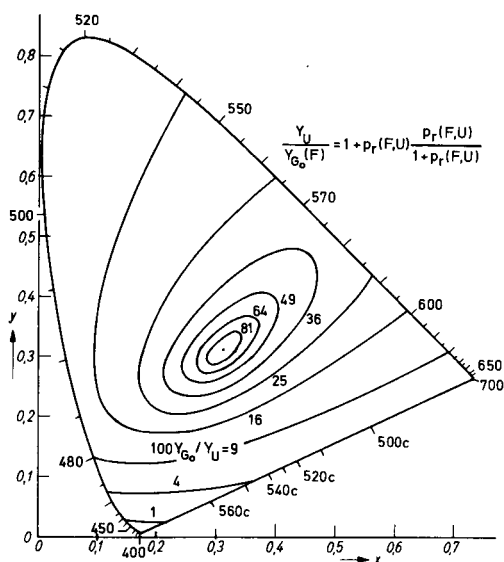


Fig. 3. Luminance ratio Y_{G_0}/Y_U of G_0 -colours in the chromaticity diagram (x, y) computed by Eq. (8).

Presently we conclude that all published experiments of Evans and Swenholt (1967, Figs 3, 6, 7, 8; 1969, Figs 5, 6, 7) can be described in good agreement by our Eqs (8) and (10) or their approximations (8a) and (10a). The following equation of these authors (Evans and Swenholt, 1969, p. 633)

$$G_0 = \log Y_U/Y_{G_0}(F) = \log [p_c(S_B - 1) + 1] \quad (11)$$

is identical with our Eq. (8a). Comparing both leads to

$$\bar{k}(\lambda, U) = (S_B - 1)/p_r(\lambda, U) \quad (12)$$

As far as we know this equation gives the first theoretical derivation of the quantity S_B (specific hue brilliance).

G_0 -COLOURS FOR CHROMATIC SURROUNDS

We define the chromatic purity $p_r(F, W)$ of Eq. (5) in agreement with Eq. (4). The chromatic purity components $p_i(W)$ are determined by the purity coordinates of the chromatic surround and CIE-illuminant C

$$p_i(W) = p_i(\text{ill. } C) + a_i (p_i(U) - p_i(\text{ill. } C)) \quad (i=2, 3) \quad (13)$$

The constants $0 \leq a_i \leq 1$ describe the amount of adaptation. We use $p_i(W) = p_i(\text{ill. } C)$ for $a_i = 0$ (no chromatic adaptation). $p_i(W) = p_i(U)$ is valid for $a_i = 1$ (complete chromatic adaptation).

Fig. 4 shows the luminance factors of G_0 -colours for different monochromatic surrounds. These luminance factors are computed by Eq. (5) with the assumption of middle adaptation. If we compare these results with the corresponding experiments of Evans and Swenholt (1969, Fig. 2), we come to the conclusion that Eq. (5) is qualitatively valid for chromatic surrounds, too. This success demonstrates that our opponent colour concept describes well the main characteristics of colour adaptation, but we can not discuss these properties in this paper.

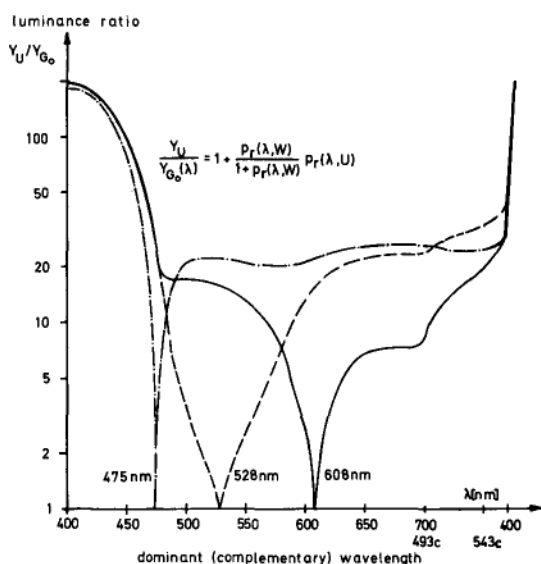


Fig. 4. Luminance ratio Y_U/Y_{G_0} of G_0 -colours for $p_c=1$ and different monochromatic surrounds ($\lambda=475$ nm; $a_i=0.25$; $\lambda=528$ nm; $a_i=0.50$; $\lambda=608$ nm; $a_i=0.50$)

C_0 -COLOURS

The C_0 -colours are determined by Evans and Swenholt (1967) for achromatic surrounds only. Their experiments lead to the following essential results.

1. The luminance factor of C_0 -colours is by a factor ten smaller than the luminance factor of the G_0 -colours for $p_c = 1$.
2. Additive mixture of any amount of achromatic colour to the spectral C_0 -colour leads to new C_0 -colours.

Looking again at Fig. 1, the C_0 -colour is for $p_c = 1$ ten times darker than the G_0 -colour. The chromatic moment M_r remains constant, if we add an achromatic colour to the C_0 -colour for $p_c = 1$. The C_0 -colours are located on parallels to the achromatic axis.

We can formulate this experimental result in the following equations

$$\frac{Y_{C_0}(\lambda)}{Y_U} = p_s \frac{Y_{G_0}(\lambda)}{Y_U} \quad (14)$$

$$\frac{M_{C_0}(F)}{M_{C_0}(\lambda)} = 1 = \frac{Y_{C_0}(F) p_r(F, U)}{Y_{C_0}(\lambda) p_r(\lambda, U)} \quad (15)$$

It follows from Eq. (8) and the approximation $\bar{k}(\lambda, U) = 1$

$$\frac{Y_{C_0}(F)}{Y_U} = \frac{p_s}{p_r(F, U)} \quad (16)$$

We compute the luminance factors of the C_0 -colours with Eq. (16) and show the results in the chromaticity diagram (x, y) in Fig. 5.

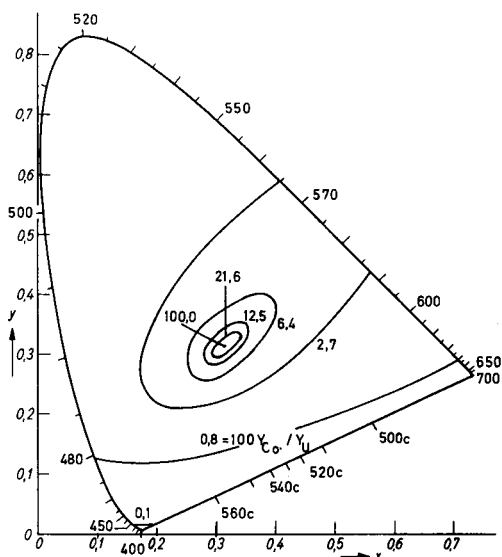


Fig. 5. Luminance ratio Y_{C_0}/Y_U of C_0 -colours in the chromaticity diagram (x, y)

Our Eq. (16) gives a good agreement with the experiments of Evans and Swenholt (1967, Fig. 8). For our example in Fig. 1 we calculate $Y_{C_0}/Y_U=0.0066$ and $M_{C_0}=0.1$. Studying Fig. 1 we notice, that the luminance factor of the added amount of spectral colour must be the same for getting a C_0 -colour on the white and black point.

So it becomes clear, why Hurvich and Jameson (1955, p. 606, Fig. 5) measure the chromatic purity $p_r(\lambda, U)$, when they project a just perceived amount of spectral colour on an achromatic field. If further the chromatic thresholds are equal at the black and white point, then the just noticeable colour differences of complementary optimal colours must be the same because of the symmetry of the optimal colour solid. Valberg and Holtsmark (1972) presented experimental evidences of such symmetries at this symposium.

L_0 -COLOURS

The L_0 -colours were measured by Kowaliski (1969). Fig. 6 shows his experimental results (x) for high purity

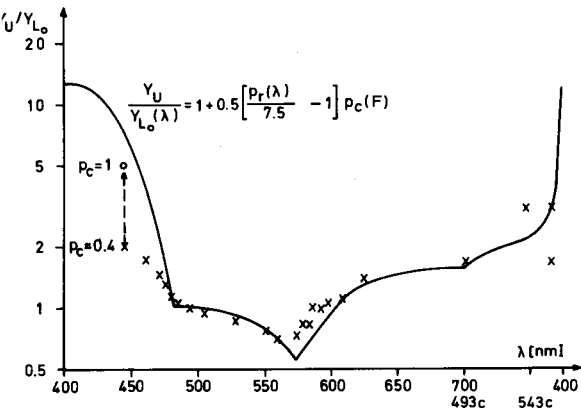


Fig. 6. Luminance ratio Y_U/Y_{L_0} of L_0 -colours for $p_c=1$ experimental (x) by Kowaliski (1969)

We describe the L_0 -colours by the following equation

$$\frac{Y_U}{Y_{L_0}(F)} - 1 = p_r(F, U) \bar{c}(\lambda, U) \tag{17}$$

with

$$\bar{c}(\lambda, U) = \frac{0.5}{p_r(\lambda, U)} \left[\frac{p_r(\lambda)}{7.5} - 1 \right] \tag{18}$$

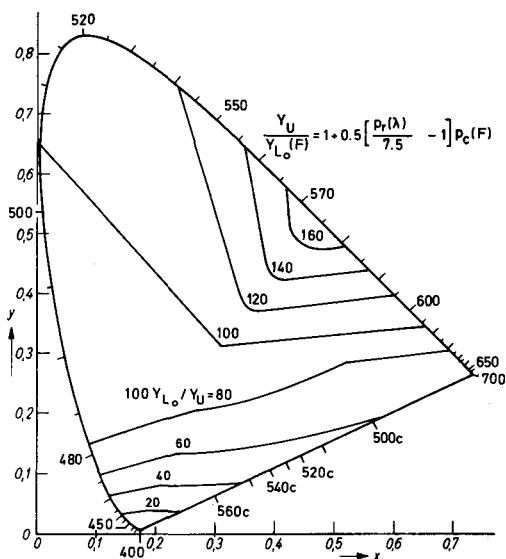


Fig. 7. Luminance ratio Y_{L0}/Y_U of L_0 -colours in the chromaticity diagram (x, y)

The results, computed by this equation in Fig. 6, describe qualitatively the experiments of Kowalski (1969). We must consider, that they were only measured up to $p_c \approx 0.4$ in the blue spectral region. If we extrapolate the experiments by Eq. (17) up to $p_c = 1$, we get a better agreement with the theory.

We calculate further the L_0 -colours of equal luminance factor in the chromaticity diagram by Eq. (17). The results may give a qualitative approximation to the experiments of Kowalski (1969) and others.

If we compare Eq. (8a) for G_0 -colours with Eq. (17) for L_0 -colours there is only an alteration of $\bar{k}(\lambda, U)$ to $\bar{c}(\lambda, U)$. This means, that we come to a new equation for L_0 -colours corresponding to equation (10a) for G_0 -colours.

$$\frac{Y_{L0}(F)}{Y_U} = 1 - \frac{\bar{c}(\lambda, U)}{Y_U} M_{L0}(F) \quad (19)$$

This equation describes the L_0 -colours in Fig. 1.

RELATIONSHIP BETWEEN MUNSELL-CHROMA, G_0 -AND L_0 -COLOURS

We have deduced for the description of Munsell-chroma (Richter, 1969c)

$$C_r = c_Y \frac{k(\lambda) p_r(F, U)}{1 + c(\lambda) p_r(F, U)} \text{ with } c_Y = 0.91 Y(F)^{0.341} \quad (20)$$

The functions $k(\lambda)$ and $c(\lambda)$ are determined by the colours of the Munsell colour system. They are shown in Fig. 2 and are tabulated in Table I of that paper. As demonstrated by the Figs 4, 6, 7, 8 and 10 of that paper, our chroma-formula describes the colour attribute Munsell chroma as well as other formulas

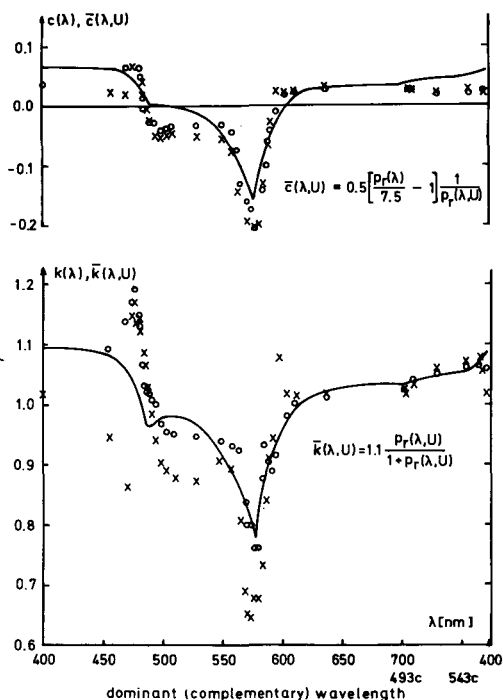


Fig. 8. Comparison of the functions $k(\lambda)$, $c(\lambda)$ (Richter, 1969c) and $k(\lambda, U)$, $c(\lambda, U)$ of Eqs [7] and [18] as function of dominant (complementary) wavelengths.

(for instance the Cube-Root formula) and can be used as colour difference formula. The discussed possibility, that we may perhaps succeed in deducing the functions $k(\lambda)$ and $c(\lambda)$ seems verified by Fig. 8. In Fig. 8 we compare the experimental results (x and o) with the functions $1.1 \bar{k}(\lambda, U)$ and $\bar{c}(\lambda, U)$ used in this paper.

The qualitative agreement of the functions $k(\lambda)$, $c(\lambda)$ with $\bar{k}(\lambda, U)$, $\bar{c}(\lambda, U)$ allows important conclusions.

On the one hand we get a theoretical chroma-formula which transforms the

$$C_r = c_Y \frac{1.1 \bar{k}(\lambda, U) p_r(F, U)}{1 + \bar{c}(\lambda, U) p_r(F, U)} \quad (21)$$

colour coordinates x , y , Y of equal Munsell chroma approximately at equally spaced circles in the whole colour solid.

On the other hand we get a theoretical chroma-formula, if we use our Eqs (8a), (17) and (21)

$$C_r = 1.1 c_Y \frac{Y_U/Y_{G_0}(F) - 1}{Y_U/Y_{L_0}(F)} \quad (22)$$

This equation leads to a surprising simple relationship between Munsell chroma, G_0 - and L_0 -colours.

The affirmation of Evans and Swenholt (1968), that colours of equal Munsell chroma correspond to an equal luminance factor of G_0 -colours is a good approximation of Eq. (22). The change of $Y_U/Y_{L_0}(F)$ is for constant chroma much smaller than the altering of $Y_U/Y_{G_0}(F)$ (compare Fig. 1). If we neglect the change of $Y_U/Y_{L_0}(F)$, the affirmation of Evans and Swenholt (1968) is evident.

CONCLUSION

Three perceived colour groups have been studied: G_0 -colours (between neither containing gray nor not containing gray), C_0 -colours (between neither achromatic nor chromatic) and L_0 -colours (equal lightness). The theoretical luminance ratio Y_F/Y_U of the central colour F and the surrounding colour U is shown in the CIE chromaticity diagram (x, y) for G_0 -, G_0 - and L_0 -colours and compared with different experimental results. The agreement between the experimental and theoretical data is good. A detailed comparison of all derived formulas leads to an analytic connection between Munsell-chroma, G_0 - and L_0 -colours. This is a proof of the relationship between different colour attributes and our colour difference formula.

REFERENCES

- Evans, R. M., and Swenholt, B. K. (1937) Chromatic strength of colors: Dominant wavelength and purity. *J. Opt. Soc. Amer.* 57, 1319-1324.
- Evans, R. M., and Swenholt, B. K. (1968) Chromatic strength of colors, Part II: The Munsell system. *J. Opt. Soc. Amer.* 58, 580-584.
- Evans, R. M. and Swenholt, B. K. (1969) Chromatic surrounds and discussion. *J. Opt. Soc. Amer.* 59, 628-634.
- Hård, A. (1969) Qualitative attributes of colour perception. Proc. 1st AIC congress "Color 69" Stockholm, 351-366.
- Hurvich, L. M., and Jameson, D. (1955) Some quantitative aspects of an opponent colors theory. I. Chromatic responses and spectral saturation. II. Brightness, saturation and hue in normal and dichromatic vision. *J. Opt. Soc. Amer.* 45, 546-552, 602-616.
- Kowaliski, P. (1969) Equivalent luminances of colors. *J. Opt. Soc. Amer.* 59, 125-130.
- Richter, K. (1969a) Antagonistische Signale beim Farbensehen und ihr Zusammenhang mit der empfindungsgemässen Farbordnung. Dissertation Universität Basel, 152.
- Richter, K. (1969b) New opponent colour concept for deriving Munsell hue and chroma as well as Evans' recent results. Proc. 1st AIC congress "Color 69" Stockholm, 403-417.
- Richter, K. (1969c) Farbdifferenzformel des Gegenfarbensystems. *Farbe* 18, 207-220.
- Richter, K. (1970) Der empfindungsgemässe Farbton als Funktion von Sättigung und Helligkeit. *Farbe* 19, 277-282.
- Valberg, A., and Holtsmark, T. (1971) Similarity between JND-curves for complementary optimal colors. This Symposium.

DISCUSSION

Valberg: From your paper I understand that you have incorporated in your model of colour appearance the experimental results obtained by Evans and

Swenholt for chromatic thresholds relative to different gray colours ranging from black to white. For these C_0 -colours the chromatic moment is constant and since in the purity diagram they are situated on a circle around the white point, the results imply a symmetry of discrimination around the $Y=50$ reflectance point of the achromatic axis (Fig. 1). Does your model also indicate symmetry properties in other directions in colour space?

K. Richter: The experimental results of Evans and Swenholt on chromatic thresholds are incorporated in the model. As you can see in Fig. 1, the colours at the chromatic threshold, called C_0 -colours, have the same chromatic moment. Therefore the chromatic threshold is the same at the black and white point in all radial directions. If you use small positive or negative slits in your JND-experiments (compare your paper), the corresponding optimal colours vary in the colour solid of Fig. 1 in angular direction. In this case the colour vectors correspond to different dominant or compensatory wavelengths. Because the space of Fig. 1 is developed as uniform near the achromatic axis, a symmetry of colour discrimination for complementary optimal colours must occur. This is true for small slits. I do not know if there is a symmetry in the case of wide slits.

Wyszecki: With regard to Fig. 7 of the preprint I would like to see plotted actual observed contour lines to compare them with your theoretical lines. I find your theoretical lines in contradiction with most, if not all, observed data. In particular, I doubt the validity of the 100-straight line to the $\lambda_d=510$ nm point.

K. Richter: Fig. 7 is based on, I hope, serious experiments of Kowaliski. Plots of his experimental results show straight lines for DIN hues 7 and 20. DIN hue 20 corresponds to $\lambda_d=510$ nm and therefore we get for $\lambda_d=510$ a straight line which is shown in my Fig. 7 (compare also Fig. 1 of Dr. Kowaliski's paper at this symposium).

Kowaliski: Dr. Richter's plot of our results is based on an extrapolation to $p_e=1.0$ which actually was never measured as all heterochromatic matches for the determination of equivalent luminances were obtained with trichromatic mixtures of $p_e < 1.0$.

K. Richter: In most of your new experiments, published at this symposium, you nearly measure up to $p_e=1$. Because many of your plots are approximately linear as function of p_e , it is easy to extrapolate up to $p_e=p_c=1$. The results are shown in my Fig. 6.

There is a discrepancy between Dr. Kowaliski's and other observed data (see also discussion of MacAdam). My feeling is, that Dr. Kowaliski's observed data may belong to colours of equal lightness but other observed data may belong more or less to colours of zero blackness (grayness), which show indeed no greater luminance ratio in the yellow region (compare my Fig. 3). So it would be useful, if Dr. Kowaliski would again interpret his experiments.

Did you measure the luminance ratio of colours of equal (perceived) lightness?

Kowaliski: Equivalent luminances show essentially the difference between the amounts of energy (flux) contained in the colored and achromatic stimuli respectively, corresponding to visual lightness or brightness matches. (Further details in my paper of this afternoon and also Kowaliski, 1969).

Billmeyer: Some years ago we studied the variation of Munsell hue with lightness at the maximum chromas available with transparent dyed acrylic plastic. In the *RP* to *R* region, the Munsell Renotation system predicts that constant-hue lines shift clockwise towards *P* as value decreases. Visual observations did not confirm this shift. Has anyone else seen this effect? (See Billmeyer *et al.*, 1961)

K. Richter: In my 1970 paper the change of hue as function of the luminance factor is described. The results give good agreement with the change of hue in the Munsell System. Some of our own experiments, published also in this paper show a shift clockwise towards *P* for unique red colours but not so large as predicted by the Munsell system.

REFERENCES

- Billmeyer, F. W., Beasley, J. K. and Sheldon, J. A. (1961). Colororder system predicting constant hue. *J. Opt. Soc. Amer.* 51, 656-666.
 Kowaliski, P. (1969) Equivalent luminances of colors. *J. Opt. Soc. Amer.* 59, 125-130.

13 PERCEPTUALLY UNIFORM SPACING OF EQUILUMINOUS COLORS AND THE LOCI OF CONSTANT HUE

DEANE B. JUDD*

*National Bureau of Standards
Washington, D.C. 20234*

INTRODUCTION

Schrödinger (1920) suggested that the most direct transition from any highly saturated color to neutral is that which passes through colors seen to have the same hue. He said (MacAdam, 1970) "It seems to me probable that the most rapid transition from any highly saturated hue to huelessness takes place without any change of hue, because such a change, as a superfluous, newly added distinguishing characteristic between neighboring colors of the series, would be liable to prolong the series."

According to Schrödinger's formulation, the shortest paths to a hueless color, that is, the color geodesics to neutral, are as shown in Fig. 1 for equiluminous colors. This formulation is based on the assumption that the red, green, and blue primaries proposed by König and Dieterici (1892) are the true fundamental colors.

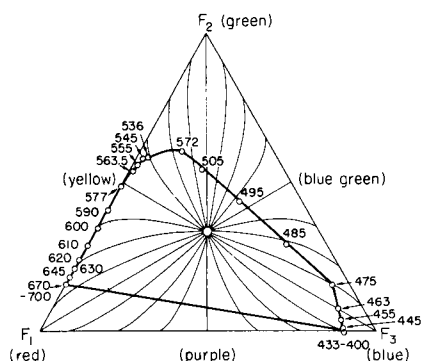


Fig. 1. Theoretical constant-hue loci (Schrödinger, 1920).

* Because of illness of dr. Judd, this paper was read by dr. I. Nimeroff.

DETERMINATIONS OF CONSTANT-HUE LOCI

In planning the Munsell color system, A. H. Munsell (Tyler and Hardy, 1940) originally intended to assign Munsell hue notations in accord with additive mixtures on a Maxwell disk. The rule was to be that when a chromatic color is mixed additively with a neutral (white, gray, or black), the hue of the mixture would be taken to be the same as that of the chromatic color. By this rule the intention was to make constant Munsell hue mean the same as constant dominant wavelength.

Observation of colors formulated to have constant dominant wavelength showed, however, that this definition of constant hue was usually unacceptable. Even the first Atlas of the Munsell Color System issued in 1915 (Gibson and Nickerson, 1940) deviated from this definition, and the color standards issued by the Munsell Color Company were progressively revised to deviate more and more from constant dominant wavelength to produce the Munsell Book of Color (Glenn and Killian, 1940; Kelly, Gibson and Nickerson, 1943). The colors of this book were then subjected to a systematic scrutiny (Newhall, 1940; Newhall, Nickerson, and Judd, 1943) that resulted in the definition of the Munsell renotations. The degree to which the present Munsell renotation hue departs from constant dominant wavelength may be seen from Fig. 2

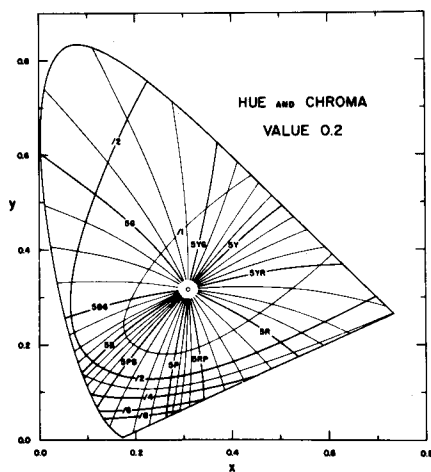


Fig. 2. Loci of constant Munsell hue and chroma.

taken from a paper by Judd and Wyszecki (1956). It will be noted that, with one exception, the lines radiating out from the center of the chromaticity diagram are somewhat curved. These lines, based on extensive observations by more than 40 observers, define constant Munsell renotation hue; straight lines from the center correspond to constant dominant wavelength. These data constitute the first results on loci of constant hue to be considered. Note that for blue the locus is curved.

A second body of data is that by MacAdam (1950) who reported for two observers the loci of constant hue determined with various surrounding colors. These loci indicated that the hueless point is controlled largely by the color of the surround, but otherwise agreed, in general, with the curvatures of constant-hue lines in the Munsell renotation system.

A third body of data is that by Wilson and Brocklebank (1955) who matched various rotary mixtures on the Maxwell disk of each of 120 relatively pure colors, forming a hue scale, with a white by having the observer point out the one of the 120 pure colors yielding a hue match. Fig. 3 shows some of their

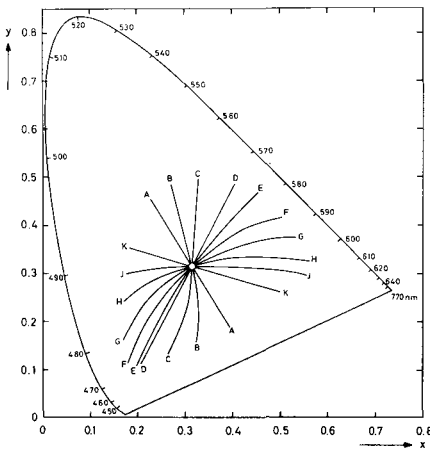


Fig. 3. Experimental constant-hue loci (Wilson and Brocklebank, 1955).

results. They agree with the Munsell and MacAdam data in general. Note, however, that the line of constant hue from the violet end of the spectrum is straight. Schrödinger made a check of this point, and remarked that the theoretical result based upon the König blue primary at 460 nm was wrong. He suggested that the choice of primary should have been violet, instead of blue, to accord with König's later studies of observers with tritanopic vision (MacAdam, 1970, p. 180). So much for data on constant-hue loci by experiment.

DETERMINATION OF COLOR GEODESICS

The task of determining the geodesic between two colors by direct experimentation is so formidable that to my knowledge no report of the results of such an experiment has ever been published. What must be done is to find a series of colors corresponding to each of a reasonable sampling of the paths in color space leading from the one color to the other, determine the number of perceptually equal steps in each path, and then pick out the path having the least number of steps. Even if we pick two colors of the same luminance

and investigate only paths involving constant luminance, the task is still discouragingly formidable. Suppose, however, you have a plane diagram whose points represent equi-luminous colors and such that you believe it has uniform chromatic spacing; then it is easy to find the shortest path implied by that belief. Simply draw a straight line, and there you have the implied geodesic.

In the Munsell color system, equiluminous colors are represented by planes whose points are identified by a polar-coordinate system with chroma corresponding to the radius vector, and hue, to angle. The spacing in such planes is intended to be uniform, and is as uniform as experiments carried out between 1915 and 1929 could make it. As we have seen, however, the colors to be represented along any radius vector were chosen to have constant hue. It was always supposed, though rarely stated in those days, that constancy of hue implied shortest path to neutral. Nickerson (1936) proposed an index of the perceptual size of a color difference. This index was called the Index of Fading. This index had the form of a sum of suitably weighted hue, chroma, and value components of the color difference: and this form of index clearly implies that, if the two colors have the same Munsell hue and value ($\Delta H = \Delta V = 0$), then constancy of hue does indeed imply the shortest path to neutral because the formula for the Index of Fading evaluates the difference as proportional solely to the chroma difference, that is to the length of a straight line, some part of the radius vector. The assumption implied by the Munsell color system is thus equivalent to Schrödinger's assumption stated backward; that is, instead of saying, as Schrödinger did, that we have an evaluation of the geodesics to neutral, and we assume that they correspond to constant hue, the Munsell system says we have determined the loci of constant hue, and we assume that they are the shortest paths to neutral.

MacAdam (1942) reported on an extensive series of determinations of the sensitivity of one observer (P. G. Nutting, Jr.) to chromaticity change at constant luminance. The standard deviation of 50 settings for a color match was used as the inverse measure of sensitivity, and the result of traversing each of 25 chromaticities in several directions (4 to 9) yielded what have become known as the "MacAdam Ellipses". MacAdam was impressed at the time by the fact that the spacing implied by these ellipses could only be represented on a plane by making adjustments somewhat larger than the estimated uncertainties. So MacAdam avoided deriving any such plane diagram for the time being. Moon and Spencer (1943) and Fry (1945), however, did not hesitate to do so; nor did Farnsworth (1957). The plane diagram derived by Fry to convert the MacAdam ellipses into equal-sized circles accorded with the modulation theory proposed by Fry. This theory is based on the opponent-color pairs, red-green and yellow-blue, as in the Hering theory, and it also accommodated the facts of both protanopia and deuteranopia. Like Schrödinger, Fry then "assumed that the lines of constant hue are straight lines radiating from white" on his

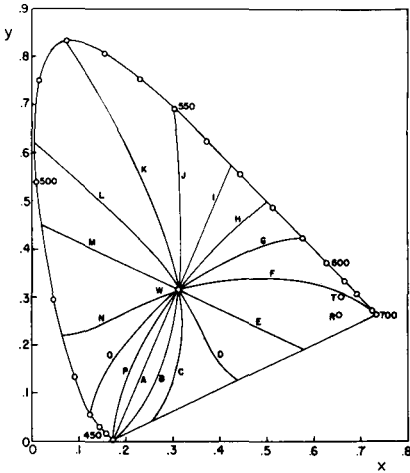


Fig. 4. Constant-hue loci derived (Fry, 1945) from MacAdam (1942) ellipses.

(u'', v'') -diagram. By reverse transformation to the (x, y) -chromaticity diagram he obtained the constant-hue loci shown in Fig. 4. Note that there are two directions in which constant-hue accords with constant dominant wavelength—one in the violet-green-yellow direction agreeing with tritanopic confusions with neutral, the other in the red-green direction agreeing with deuteranopic confusions with neutral.

After a delay of 27 years, MacAdam (1969) derived his own nonlinear transform of the (x, y) -diagram adjusted to reduce as closely as possible the MacAdam ellipses to equal-size circles. This transform was produced by plotting in rectangular coordinates two variables, XI and ETA. XI and ETA are ten-term polynomials in auxiliary variables, a and b , which are, in turn, projective transforms of the chromaticity coordinates, x, y . The 26 constants involved in the double transformation were optimized by automatic digital computer. Fig. 5 shows the curved lines in the (x, y) -chromaticity diagram that correspond

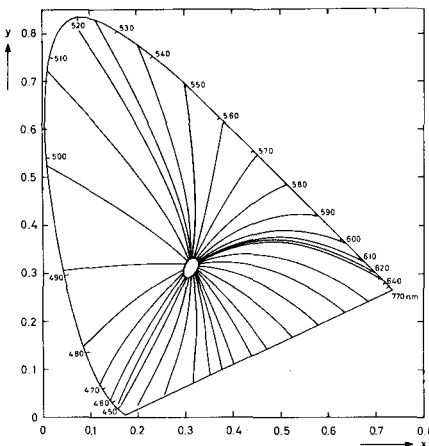


Fig. 5. Constant-hue loci derived (MacAdam, 1969) from the MacAdam (1942) ellipses.

to straight lines drawn on the (XI-ETA)-chromaticity diagram intended to have uniform spacing. In accord with the usual assumption MacAdam has labeled these curved lines "constant-hue loci". Note that this strictly empirical finding from the MacAdam ellipses indicates that the line of constant dominant wavelength to the violet extreme of the spectrum is very closely the shortest path; but for blue and for reddish violet the shortest paths are curved if plotted on a mixture diagram.

Still another determination of the geodesics implied by the MacAdam ellipses was recently reported by Muth and Persels (1971). He started with the formula for the perceived size of color differences derived from the MacAdam ellipses by Friele (1965, 1966) and refined by MacAdam (1966) and Chickering (1967). By computer he determined the tri-dimensional geodesics required to pass from various of the spectrum colors to a white of unit luminance. The relative luminance of the spectrum color was adjusted until the geodesic approached white with a zero slope. The vertical projections of these tridimensional geodesics onto the (x,y)-chromaticity diagram are shown by courtesy of Persels in Fig. 6.

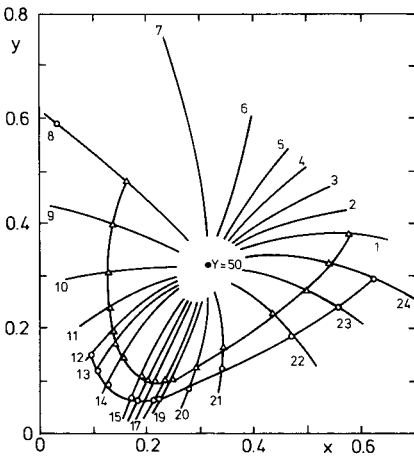


Fig. 6. Vertical projections of Persel's three-dimensional constant-hue loci derived from the Friele (1965, 1966) and MacAdam (1966) formulas with constants optimized by Chickering (1967). (from Muth and Persel, 1971).

This is the last derivation of the geodesics implied by the MacAdam ellipses to be considered. The shapes of these geodesics plotted on the (x,y)-chromaticity diagram agree well.

A method independent of the MacAdam ellipses has been used by Yonemura (1970) to derive geodesics to neutral. This method is based on the opponent-colors theory as exemplified by the CIE 1960 UCS diagram originally derived by MacAdam (1937) some years before he developed the MacAdam ellipses. This diagram has been shown by Judd and Yonemura (1970) to accord with the second stage of the Müller theory. It is a projective transformation of the (x,y)-

chromaticity diagram. Yonemura made use of measures of red-green and violet-greenyellow activity that took the form of angles subtended by the chromaticity-difference vector at the tritanopic convergence point, and at the deuteranopic convergence point, respectively. The argument is that normal vision may be viewed as a combination of tritanopic and deuteranopic vision as originally suggested by Walraven and Bouman (1965). Since the chromaticities along any line passing through the tritanopic convergence point cannot be distinguished by the tritanope, the angle measure of red-green distinction made by him is mandatory. By this argument the diagram produced by plotting in rectangular coordinates the angles subtended at the deuteranopic convergence point against those at the tritanopic convergence point is a uniformly spaced chromaticity diagram. It is, of course, a nonlinear transform of the (x,y) -chromaticity diagram. If the spacing of this nonlinear transform is accepted as a useful approximation to uniform spacing, then the series of equiluminous colors defined by any straight line on this diagram is a useful approximation to the geodesic between the colors represented by the extremes of the line. Neutral colors are represented at the origin of the coordinate system formed by plotting θ_{vgy} against θ_{rg} . Any straight line from this origin in a direction between red and greenyellow not only corresponds to a geodesic but also to colors for which the ratio of red to greenyellow is constant. But this is the condition for constancy of hue. So by this theoretical construct, no separate assumption that a color geodesic corresponds to a locus of constant hue is required; the construct already implies this correspondence. Fig. 7 is taken from Yonemura (1970) and compares the geodesics to neutral derived from his angle version of the CIE 1960 UCS diagram with some of the loci of constant hue according to the Munsell renotations. Note that the agreement is good even in the violet

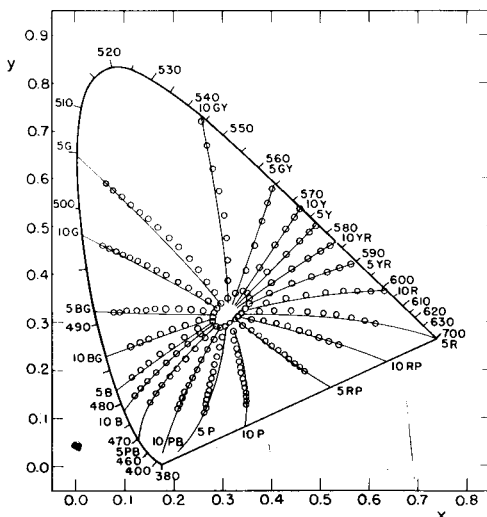


Fig. 7. Constant-hue loci (Yonemura, 1970) compared to Munsell-renotation-hue loci.

to reddish violet hue region where the theoretical prediction accords better with the MacAdam and the Wilson-Brocklebank data than with the Munsell renotations. Note that the constant-hue loci are straight for only four directions, those toward the tritanopic convergence point and its complementary, and those toward the deuteranopic and its complementary.

The second theoretical construct to be considered is essentially that of the stage theory of color vision proposed by Adams in 1923, and elaborated by him in 1941. The measure of red-green activity in the third stage is taken as $V_x - V_y$, and that of yellow-blue activity as $V_y - V_z$, where V_i stands for the Munsell value function of tristimulus value T . Adams recognized that this theoretical construct failed to agree with the Munsell constant-hue loci in the blue to violet region because the predicted locus for blue is straight rather than curved, but he suggested that further experiments might support the theory rather than the Munsell renotations.

Glasser, Reilly, et al (1958) have for some years been exploring the adaptability of the Adams chromatic-value diagram to predictions of the perceived sizes of color differences. They used the cube-root function instead of the Munsell-value function, and they did not restrict it to the XYZ primaries of the 1931 CIE coordinate system as Adams did, but chose three other primaries RGV adjusted to improve the predictions of perceived size of color differences. In a discussion of these choices with Reilly in June 1970, it became clear that choice of R , G , and V to correspond to the convergence points of the three types of dichromatic vision, as suggested by Schrödinger in 1920, would make the Adams chromatic-value diagram yield much improved predictions of the constant-hue loci. I have worked out the constant-hue loci implied by it. The hue index used was $(R^{1/3} - G^{1/3})/(G^{1/3} - V^{1/3})$. Note that if $R = G$, this index remains constant at zero regardless of the values of R , G , and V , and if $G = V$, it is constant at infinity. This accounts for the straight lines in the directions

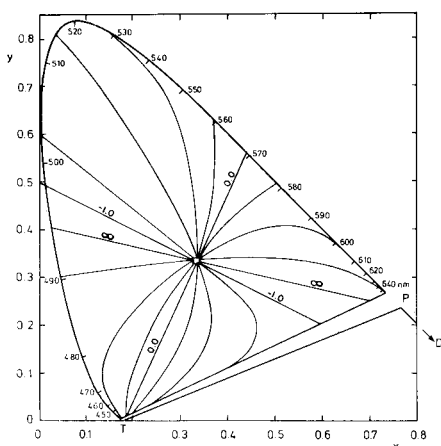


Fig. 8. Cube-root constant-hue loci. Straight constant-hue loci pass through dichromatic convergence points P, D and T.

of the tritanopic and protanopic convergence points and their complementaries. There is also a straight line in the direction of the deuteranopic convergence point and the opposite direction toward its complementary. All points on this line correspond to $R=V$, and for such points the hue index becomes: $(1 - G^{1/3}) / (G^{1/3} - 1) = -1$, regardless of the value of $G^{1/3}$. Fig. 8 shows the lines of constant hue implied by this variation of the Adams chromatic-value diagram. The main features of the experimental loci are duplicated, but the predicted curvatures are somewhat greater than those observed. This discrepancy indicates that a power intermediate to $1/3$ and 1 might be found that would make predictions in accord with the experimental facts.

CONCLUSIONS

The assumption, made by Schrödinger (1920) that geodesics to neutral are also loci of constant hue, seems to have been amply supported by experiment. Two theoretical constructs yielding fairly successful predictions of constant-hue loci, yield this assumption as a part of their prediction.

The success of the Yonemura (1970) angle version of the CIE 1960 UCS diagram supports the suggestion that normal vision may usefully be described as a combination of tritanopic with deuteranopic discriminations as suggested by Walraven and Bouman (1965).

The limited degree of success of the Adams chromatic-value diagram, applied to the R , G , and V primaries explaining the three types of dichromatic vision as "loss" variations of normal vision still gives significant support to the view that the deviations of the loci of constant hue from constant dominant wavelength arise simply from a nonlinear connection between stimulus and response; that is, a response compression.

By this theoretical view, any direction from the neutral point yielding a straight line on the mixture diagram as a locus of constant hue must be toward the point representing a primary color process.

There are thus many more or less equally successful ways to explain color geodesics and the loci of constant hue. Psychophysics has thus indicated a number of explanations, but it could easily take many years of physiological study to decide which, if any, of these explanations is correct.

REFERENCES

- Adams, E. Q. (1923), A theory of color vision. *Psychol. Rev.* 30, 56-76
- Adams, E. Q. (1942), X-Z planes in the 1931 ICI system of colorimetry. *J. Opt. Soc. Amer.* 32, 168-173
- Chickering, K. D. (1967), Optimization of the MacAdam-modified 1965 Friele color-difference formula. *J. Opt. Soc. Amer.* 57, 537-541
- Farnsworth, D. (1957), A temporal factor in colour discrimination. In: "Visual Problems

- of Colour", Symposium September 1957 in Teddington, Vol. 2, 429-444
- Friele, L. F. C. (1965), Further analysis of color discrimination data. *J. Opt. Soc. Amer.* 55, 1314-1319
- Friele, L. F. C. (1966), Friele approximations for color metric coefficients. *J. Opt. Soc. Amer.* 56, 259-260
- Fry, G. A. (1945), A photo-receptor mechanism for the modulation theory of color vision. *J. Opt. Soc. Amer.* 35, 114-135
- Gibson, K. S., and Nickerson, D. (1940), An analysis of the Munsell color system based on measurements made in 1919 and 1926. *J. Opt. Soc. Amer.* 30, 591-608
- Glasser, L. G., McKinney, A. H., Reilly, C. D. and Schnelle, P. D. (1958), Cube-root color coordinate system. *J. Opt. Soc. Amer.* 48, 736-740
- Glenn, J. J., and Killian, J. T. (1940), Trichromatic analysis of the Munsell Book of Color. *J. Opt. Soc. Amer.* 30, 609-616
- Judd, D. B., and Wyszecki, G. (1956), Extension of the Munsell renotation system to very dark colors. *J. Opt. Soc. Amer.* 46, 281-284
- Judd, D. B., and Yonemura, G. T. (1970), CIE 1960 UCS diagram and the Müller theory of color vision. *J. Research National Bureau of Standards* 74A, 23-30
- Kelly, K. L., Gibson, K. S., and Nickerson, D. (1943), Tristimulus specification of the Munsell Book of Color from spectrophotometric measurements. *J. Opt. Soc. Amer.* 33, 355-376
- König, A., and Dieterici, C. (1892), Die Grundempfindungen in normalen und anomalen Farbensystemen und ihre Intensitäts-verteilung im Spectrum. *Z. Psychol. u. Physiol. der Sinnesorgane* 4, 241-269
- MacAdam, D. L. (1937), Projective transformations of the I.C.I. color specifications. *J. Opt. Soc. Amer.* 27, 294-299
- MacAdam, D. L. (1942), Visual sensitivities to color differences in daylight. *J. Opt. Soc. Amer.* 32, 247-274
- MacAdam, D. L. (1950), Loci of constant hue and brightness determined with various surrounding colors. *J. Opt. Soc. Amer.* 40, 589-595
- MacAdam, D. L. (1965), Analytical approximation for color metric coefficients IV. Smoothed modifications of Friele's formulas. *J. Opt. Soc. Amer.* 55, 91-95
- MacAdam, D. L. (1966), Smoothed versions for Friele's approximation for color metric coefficients. *J. Opt. Soc. Amer.* 56, 1784-1785
- MacAdam, D. L. (1969), Geodesic chromaticity diagram. *Proc. 1st AIC Congress, "Color 69"*, Stockholm; 293-301. *Farbe* 18, 77-84
- MacAdam, D. L. (1970), *Sources of Color Science*, Cambridge, Mass., MIT pr., see p. 178 for his translation.
- Moon, P., and Spencer, D. E. (1943), A metric for colorspace. *J. Opt. Soc. Amer.* 33, 260-269
- Muth, E. J., and Persels, C. G. (1971), Constant-brightness surfaces generated by several color-difference formulas. *J. Opt. soc. Amer.* 61, 1152-1154
- Newhall, S. M. (1940), Preliminary report of the O.S.A. Subcommittee on the Spacing of the Munsell Colors. *J. Opt. Soc. Amer.* 30, 617-645
- Newhall, S. M., Nickerson, D., and Judd, D. B. (1943), Final report of the O.S.A. Subcommittee on the Spacing of the Munsell Colors. *J. Opt. Soc. Amer.* 33, 385-418
- Nickerson, D. (1936), The specification of color tolerances. *Textile Res.* 6, 505-514
- Schrödinger, E. (1920), Grundlinien einer Theorie der Farbenmetrik im Tagessehen. *Ann. Physik*, 63, 397-515
- Tyler, J. E., and Hardy, A. C. (1940), An analysis of the original Munsell color system. *J. Opt. Soc. Amer.* 30, 587-590
- Walraven, P. L., and Bouman, M. A. (1965), Fluctuation theory of colour discrimination of normal trichromats. *Vision Res.* 6, 567-586
- Wilson, M. H., and Brocklebank R. W. (1955), Complementary hues of after-images. *J. Opt. Soc. Amer.* 45, 293-299.
- Yonemura, G. (1970), Opponent-color-theory treatment of the CIE 1960 (u,v) diagram: Chromaticness difference and constant-hue loci, *J. Opt. Soc. Amer.* 60, 1407-1409

DISCUSSION*

Guth: In the preliminary theory which I outlined in my paper, curvilinear hue lines are caused by the fact that the *BvsY* system inhibits the *RvsG* system less and less as a stimulus is desaturated. How do the theories which you discussed cause the curvilinear loci?

Wyszecki: At the end of dr. Judd's paper you will find an answer, in that he clearly thinks in terms of a non-linear relation between stimulus and response.

Guth: That would mean that the slope of the response vs. intensity function is greater for *B—Y* than for *R—G*?

Wyszecki: I would say: that follows naturally.

Guth: It can be seen, then that for both the Bezold-Brücke effect and the curvilinear hue lines (the Abney effect) there are two equally-tenable explanations. One hypothesizes that there are different intensity-response functions, whereas the other uses a between mechanism inhibitory effect.

Nimeroff: It would take further investigation to determine which of these two hypotheses best fit the available data. We may not be able to resolve this problem even with additional information.

I would have expected that under your hypothesis the constant lines from a suitable white point would be straight lines.

Wyszecki: 1. A recent experimental determination of constant hue lines by Robertson (Stockholm meeting) may be worth mentioning.

2. It is of interest to note that many models of color vision (line elements, etc.) predict remarkably well the observed hue lines despite the fact that these models are based on fundamentally different principles of construction.

Judd: My omission of a reference to the work of Robertson was inadvertent, and I am glad to follow Dr. Wyszecki's suggestion that it be included. (Robertson, 1969).

Richter: I would like to add that in my own opponent colour model I can compute the hue lines of the Munsell system in good agreement. I get two straight lines in the yellow region for 10 *Y* and in the purple blue. This description is better than by the many other theories, which get straight hue lines in the chromaticity diagram for 3 up to 6 directions. Also my model is able to describe the change of hue with luminance factor in good agreement with the Munsell hue curves for this case (Richter, 1970).

Boynton: Recent work by Savoie and Cornsweet, (1971) reported in Sarasota has shown that the wavelengths usually thought to yield invariant hue with changes in intensity, do not actually do so. The older work by Purdy was based on only two levels (100 and 1000 trolands). Between these levels, Savoie and Cornsweet found hue changes.

* The discussion was supplemented with dr. Judd's written comments.

Vos: It should be added that Savoie and Cornsweet worked with 5 msec flashes, that is in a quite different situation from that of Purdy, e.g.

Nimeroff: It may be that luminance is the reason that the several investigations arrive at somewhat different sets of constant-hue lines.

Hard: Could constant hue-determination result in different experimental data if we look for a constant ratio between e.g. red and yellow in a simultaneous comparison situation of two samples, or if we try, by absolute judgment of the ratio, to look at the two samples each at a time without comparison, like in the NCS-experiments?

Nimeroff: If the observers can more readily make the constant-hue judgment, when the latter question is put to them than when the former question is put to them, then better, more consistent results may be obtained.

Wright: I find it of great interest that although Dr. Judd's paper deals with uniform spacing of equi-luminous colours, yet he has illustrated his argument with the non-uniform (x,y) -chromaticity chart and not with the so-called uniform (u,v) chart. I personally applaud this as the (u,v) chart seems to me an unsatisfactory chart to illustrate general chromaticity relations. The feature which makes the (u,v) chart unsatisfactory is the compression of the important yellow-green, yellow, orange area of the chart. This is brought out by the distribution of the samples on the (x,y) and (u,v) charts prepared by one of our students, Miss P. Dunster, which I will put on display.

McLaren: The Adams-Nickerson UCS diagram has been found to be free from the compressions which spoil both the (x,y) and (u,v) diagram; in addition Munsell samples of constant hue plot as straight lines.

Judd: Like Dr. Wright I have seen demonstrations that the (u,v) -chromaticity diagram is unduly compressed along the white to yellow locus; but there are conditions for which the (u,v) spacing is correct. The Priest-Brickwedde data (1926, 1938) on minimum perceptible colorimetric purity corroborate it perfectly, and Dr. Wright's own data (1943) indicate that the spacing from white to yellow for his conditions of observation should be even more compressed than is shown on the (u,v) -diagram. Use of high luminance and large angular subtense of target, and surrounds of the same high luminance open up more white-yellow discriminations than they do red-green discriminations for the same range of colours. I think that the (u,v) -spacing is much more nearly correct than the (x,y) -spacing for average viewing conditions.

REFERENCES

- Priest, I. C. and F. G. Brickwedde (1926, 1938), The minimum perceptible colorimetric purity as a function of dominant wavelength with sunlight as neutral standard. *J. Opt. Soc. Amer.* 13, 306; 28, 133-137
 Richter, K. (1970), Der empfindungsgemäße Farbton als Funktion von Sättigung und Helligkeit. *Farbe* 19, 277-282

- Robertson, A. R. (1969), A new determination of lines of constant hue, Proc. 1st AIC congress "Color 69", Stockholm, 395-402
- Savoie, R. E., and Cornsweet, T. N. (1971), The Bezold-Brücke effect and non-linearity in the visual system. Paper presented at the ARVO-meeting, Sarasota, April 30
- Wright, W. D. (1943), The graphical representation of small color differences, J. Opt. Soc. Amer. 33, 632-636

14 ROLE OF LUMINANCE INCREMENTS IN SMALL COLOR DIFFERENCES

DAVID L. MACADAM

*Research Laboratories, Eastman Kodak Company,
Rochester, N.Y. USA*

Three of the four color-difference formulas recommended for trial by the CIE (Wyszecki, 1968) imply that, for a given luminous reflectance difference, the color difference is greater when the chromaticities (x, y) of the specimens are identical than when the lighter specimen is somewhat less saturated (i.e., closer to a gray of equal reflectance). Schultze and Gall (1969), who were the first to call attention to this, found that all except the modified Friele formula define equal-color-difference ellipsoids whose tops are tipped towards white in x, y, Y space.

OBSERVATIONS CONCERNING TILTS OF ELLIPSOIDS

To test the appropriateness of such tilts, Schultze and Gall prepared and examined visually two sets of specimens, red and blue. From reports of judgments by 25 observers, they concluded that "only 3 color difference formulae gave results that were consistent with the visual assessments: the Simon-Goodwin, MacAdam-Friele-Chickering, and the DIN formulae. The agreement of the other formulae was poor or very bad, and the CIE 1964 formula gave the worst agreement by far." Of the formulas mentioned favorably, by Schultze and Gall, only the MacAdam-Friele-Chickering was among those suggested by the CIE for trials. It was derived from the same observer data as was the Simon-Goodwin formula and was intended as a replacement for that graphical method, better adapted for modern computing machines.

Further evidence in support of the conclusion of Schultze and Gall, that the ellipsoids of constant color difference are not significantly tipped, is shown in Fig. 1, after Brown and MacAdam (1949).

The directions and displacements (enlarged 75 times) of the points of highest luminance on 37 average ellipsoids for WRJB and DLM, who both observed

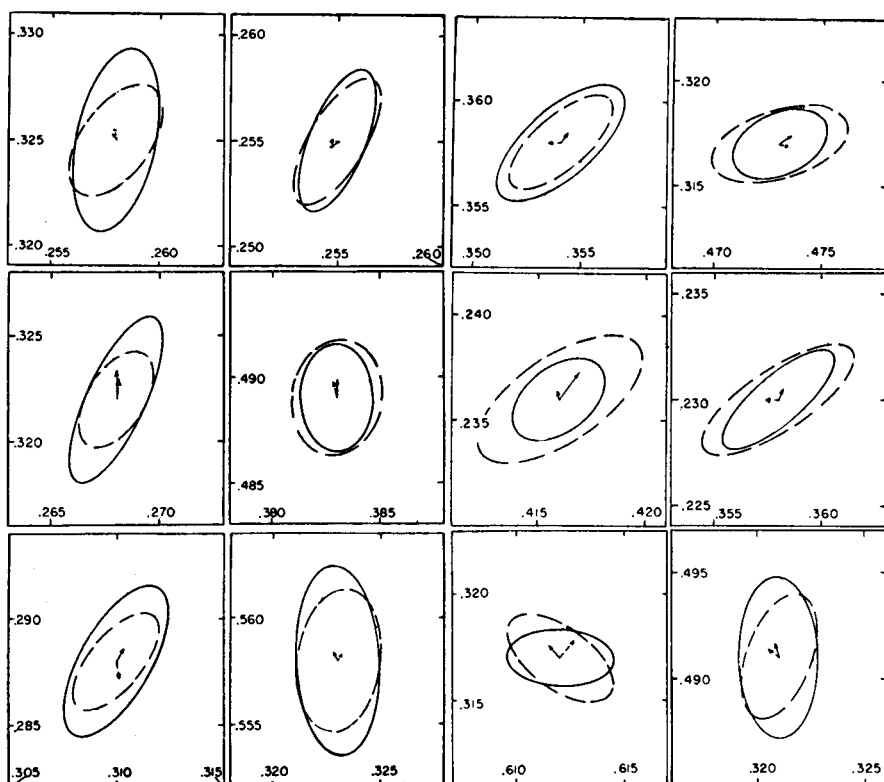


Fig. 1. Chromaticities of points of highest luminance on ellipsoids of variance of color matching (arrow heads), shown relative to constant luminance cross sections. Aim colors and centers of ellipsoids are at origins of arrows. Solid ellipses and arrows, observer WRJB; broken, DLM.

2° fields, monocularly, in a dark surround, are shown in Fig. 2. The displacements of the chromaticities of the highest points on 22 average ellipsoids for 12 young normal observers, who observed 10° fields, binocularly, in a light surround, are shown in Fig. 3. The smallness of the arrows shown in Fig. 1, relative to the corresponding radii of the ellipses, similar indications in diagrams for 23 other aim colors, shown by Brown and MacAdam (1949), and the apparent randomness of the directions of the arrows, shown most clearly in Figs. 2 and 3, indicate that, regardless of the size of the field or lightness of the surround, the tipping of the ellipsoids in x, y, Y space is not systematic or significant in comparison with the uncertainties of the observations.

CONSTANT-BRIGHTNESS GEODESICS

Persels (1971) used dynamic programming to extend toward the spectrum and several purples some geodesics (i.e., paths of least number of units of color

$Y=7$ and all of the extreme purples have Y less than 35. The points at which the geodesics fall to $Y=45$ are marked with x . Those at which $Y=40$ are marked with o . The values of Y on the geodesics at various locations in the chromaticity (x,y) diagram are very similar to the luminances necessary to match the brightness of 50 units luminance of white. This indicates that the original modified Friele formula (FMC-1) has noneuclidean properties requisite, according to the reasoning of Schrödinger (1920), for generating a realistic constant-brightness surface. As anticipated by Schrödinger, the chromaticities indicated by the geodesics also closely resemble constant-hue loci that have been inferred from appropriate series of observations.

Persels also tried using the version (FMC-2) of the modified Friele formula that was recommended for trial by the CIE (Wyszecki, 1968), but he found that the geodesics obtained with it, which are horizontal at white, C , rise to high values of Y as they approach the spectrum or purple. From this result, which contradicts experimental determinations of luminances required for equal brightness, I conclude that the noneuclidean properties of the FMC-2 formula are not consistent with experimental data on constant brightness, nor with the Schrödinger (1920) concept of brightness.

Because the U^* , V^* , W^* space of the CIE 1964 formula (Wyszecki, 1968) is euclidean, the geodesics of that space that are horizontal at white remain horizontal everywhere, implying that equal luminance is required for equal brightness for all chromaticities. The Munsell Renotation formula (Wyszecki, 1968) yields the same result, for the same reason. Persels (1971), using his dynamic programming, found that the cube-root formula (Wyszecki, 1968) generates an equal-brightness surface that is very nearly flat in x,y,Y space. The horizontal-at-white geodesic that starts at $Y=50$ falls only to $Y=49.8$ at the purple $x=0.5$, $y=0.15$; the similar geodesic to 540 nm remains at $Y=50$ over its entire length (Figure 4 indicates that the FMC-1 formula has it falling to $Y=45.3$); the geodesic to 400 nm falls only to $Y=48.2$, which may be compared with $Y=7$ for the geodesic derived from the FMC-1 formula. From these results, obtained by Persels (1971), I conclude that none of the formulas recommended for test by the CIE (Wyszecki, 1968) define color space in such a way that realistic surfaces of constant brightness can be derived by use of the Schrödinger (1920) concept of brightness.

HORIZONTAL-AT-WHITE CRITERION FOR CONSTANT-BRIGHTNESS LOCI

Because Schrödinger did not discuss colors in terms of a space equivalent to x,y,Y , some justification should be given for the decision to make the constant-brightness geodesics horizontal at white in that space. The color vectors that Schrödinger (1920) discussed on page 487 (p. 189 of the translation), which all intersect at the origin (of tristimulus space, equivalent to X,Y,Z space) repre-

sent series of colors of constant chromaticity (x, y) and correspond to parallel, vertical lines in x, y, Y space. Therefore, if the ellipsoid of equal color distance that Schrödinger (1920) discusses and shows in his Fig. 1 is symmetrical about the plane of constant Y , which is the case for the ellipsoid centered on white or any gray, according to observations, and according to all of the color-difference formulas under test, then the points where such lines are tangent to the ellipsoid all lie on the plane of constant Y through the center of the ellipsoid. Therefore, the geodesics of constant brightness should begin horizontally at white (or gray), according to Schrödinger's concept.

MODIFIED FRIELE FORMULA

Friele's (1961) derivation of his formula from a rather simple and plausible theory may be recalled. My modification merely smoothed the transitions between the different idealized cases that Friele considered. The success of the final outcome, displayed in Fig. 4, indicates that the Friele formula has the correct form and that the optimized parameters are essentially correct, despite limited precision of the observational data. In my opinion the FMC-1 formula is superior to any of the formulas recommended for test by the CIE. None of them was based on as many, or more-precise data. One of the other formulas under test yields straight lines on the x, y diagram as constant-hue loci, and equal Y as the condition for equal brightness of different chromaticities. The second yields nearly constant- Y surfaces of constant brightness. The third, besides indicating constant- Y surfaces of constant brightness, yields again only the constant-hue loci that were used, rather than direct data on color differences, in its construction. The fourth is an unfortunately distorted version of the FMC-1 formula.

$$\begin{aligned}\text{Using } P &= 0.53X + 0.28Y - 0.072Z \\ Q &= -0.35X + Y + 0.093Z \\ S &= 1.528Z\end{aligned}$$

we can write the FMC-1 measure of color difference more simply, in terms of redefined ΔL , ΔC_{rg} and ΔC_{yb}

$$\Delta E = \sqrt{(\Delta C_{rg})^2 + (\Delta C_{yb})^2 + (\Delta L)^2},$$

$$\text{where } (\Delta C_{rg})^2 = \left\{ 240 \frac{(Q\Delta P - P\Delta Q)}{(P^2 + Q^2)} \right\}^2 [1 + 2.73P^2Q^2/(P^4 + Q^4)],$$

$$(\Delta C_{yb})^2 = \left\{ 57 \left[\frac{(P\Delta P + Q\Delta Q)}{(P^2 + Q^2)} - \frac{\Delta S}{S} \right] \right\}^2 / (1 + Y^2/S^2),$$

$$(\Delta L)^2 = \left\{ 67 \frac{(P\Delta P + Q\Delta Q)}{(P^2 + Q^2)} \right\}^2 [1 + 2.73P^2Q^2/(P^4 + Q^4)].$$

The relation of this formula to the opponent-colors theory becomes clear when the case $P \simeq Q$ is examined. Then

$$\Delta C_{rg} \simeq 185 \left(\frac{\Delta P}{P} - \frac{\Delta Q}{Q} \right),$$

$$\Delta C_{yb} \simeq 40 \left(\frac{\Delta P}{2P} + \frac{\Delta Q}{2Q} - \frac{\Delta S}{S} \right),$$

and
$$\Delta L \simeq 104 \left(\frac{\Delta P}{2P} + \frac{\Delta Q}{2Q} \right).$$

The coefficient 40 in ΔC_{yb} is for $S \simeq Y$. For $S \gg Y$, that coefficient becomes 57, and for $S \ll Y$ it becomes $57 S/Y$. The appearance of $\frac{\Delta P}{P}$, $\frac{\Delta Q}{Q}$ and $\frac{\Delta S}{S}$ in these

approximations suggests that something like logarithms of the initial excitations of the retina are formed, by some nonlinear process, before the signals are opposed to form the opponent-colors responses, e.g.,

$$C_{rg} = \mathcal{L}(\log P - \log Q),$$

$$C_{yb} = \mathcal{L}(\log P + \log Q - 2 \log S),$$

and
$$L = \mathcal{L}(\log P + \log Q),$$

where $\mathcal{L}$ indicates a linear relation (proportionality plus some constant). Recall that P represents a response (a tristimulus value) weighted most heavily in the longer ("red") wavelengths, Q a response weighted most heavily in the middle ("green") wavelengths, and S a response to the shorter ("blue") wavelengths. The combination $\frac{P\Delta P + Q\Delta Q}{(P^2 + Q^2)}$ that appears in both ΔC_{yb} and ΔL and

which reduces to $\frac{\Delta P}{2P} + \frac{\Delta Q}{2Q}$ when $P \simeq Q$, and to $\log P + \log Q$ in C_{yb} and L , plays

the role of a response to yellow in ΔC_{yb} and C_{yb} itself.

In all considerations of the Friele formula, care should be taken not to confuse L with brightness, or lightness or luminance. It is none of these. The term brightness should be confined to the concept discussed by Schrödinger. Lightness is closely related to brightness, in connection with reflection colors.

DEPENDENCE OF COLOR-DIFFERENCE TOLERANCES ON LUMINANCE

The most serious objection that has been raised to the modified Friele formula (FMC-I) is that a given chromaticity difference is evaluated nearly the same whether the specimens are light or dark. Under some quite frequently encountered conditions, such as observations of opaque, reflecting specimens, a given chromaticity difference is much more noticeable with light

specimens than with dark. This was the reason for the introduction of the factor K_1 (a function of Y) in the FMC-2 formula (Chickering, 1971). A similar effect, applying to the noticeability of luminous-reflectance increments ΔY , was the reason for inclusion of K_2 in the FMC-2 formula. However, because the resulting constant-brightness geodesics are not realistic, inclusion of K_1 and K_2 destroys the unique merit of the modified Friele formula for generating the Schrödinger equal-brightness surface.

The factors K_1 and K_2 can be introduced, to get the same results as with the formula (FMC-2) recommended for test by the CIE by merely multiplying the first two coefficients, 240 and 57, by K_1 and the third, 67 by K_2 . However, in my opinion, it would be sufficient, instead, merely to change the third coefficient, 67, to 30 and to multiply the resulting value of ΔE by $(1/2)Y^{1/3}$, to provide for the greater noticeability of color differences of light than dark reflection colors. The reduction of importance of lightness increments that Strocka (1971) reported to be advantageous can be accomplished by further reducing the coefficient 67, to 20 or even 15. However, I think that such reduction of the importance of the lightness components of color differences must correspond to comparisons either of separated samples or coarse-texture materials or samples that have poorly defined or somewhat shaded (folded?) edges. In any case, the coefficient 67 can be adjusted by trial to adapt the formula for any particular observing condition. The value 67 was derived from observations in closely contiguous semicircular fields that subtended 2° , in a large surrounding field, the luminance of which was half that of the compared colors. Because these conditions were nearly ideal, the coefficient 67 is probably nearly the maximum that will ever be appropriate for any observing conditions.

The general formulas shown above for P , Q , S , $(\Delta C_{rg})^2$, $(\Delta C_{yb})^2$ and $(\Delta L)^2$ contain only 10 coefficients. Those coefficients, which are essentially the same as those that were recommended for test by the CIE, were optimized by Chickering on the basis of standard deviations of color matching by the observer PGN, whose observations were completed over 30 years ago. The 10 coefficients of the modified Friele formula can easily be re-optimized by use of any new body of color-difference data.

I am confident that such a re-optimized, modified Friele formula will still retain the desirable features that I have described, and that it will be convenient and useful for color-difference and color-tolerance applications. The final formula will be easily adjustable for specific observing conditions and applications, in the manner I explained.

REFERENCES

- Brown, W. R. J. and MacAdam, D. L. (1949) Visual sensitivities to combined chromaticity and luminance differences. *J. Opt. Soc. Amer.*, 39, 808-834

- Brown, W. R. J. (1957) Color discrimination of twelve observers. *J. Opt. Soc. Amer.*, 47, 137-143
- Chickering, K. D. (1967) Optimization of the MacAdam-modified 1965 Friele color-difference formula. *J. Opt. Soc. Amer.*, 57, 537-547. The convenient nomenclature originated by McLaren (1969) will be employed to distinguish Chickering's original form (FMC-1) of the modified Friele formula from that (FMC-2) recommended for test by the CIE (Wyszecki, 1968).
- Chickering, K. D. (1971) FMC color-difference formulas: clarification concerning usage. *J. Opt. Soc. Amer.*, 61, 118-122
- Friele, L. F. C. (1961), Analysis of the Brown and Brown-MacAdam color discrimination data. *Farbe* 10, 193-224
- McLaren, K. (1969), Scaling factors in color difference formulas: a confusing situation. *Color Eng.*, 7, 6: 38-45
- Persels, C. G. (1971), Dissertation, University of Florida; see also Muth, E. J. and Persels, C. G. (1971), Constant-brightness surfaces generated by several color-difference formulas. *J. Opt. Soc. Amer.*, 61, 1152-1154
- Schrödinger, E. (1920), *Grundlinien einer Theorie der Farbenmetrik in Tagessehen. II Teil: Höhere Farbenmetrik.* *Annalen der Physik* 63, 481-520. English translation in *Sources of color science*, Cambridge, Mass., MIT Press, 1970, pp. 155-182
- Schultze, W., and Gall, L. (1969), Experimentelle Überprüfung mehrerer Farbabstandsformeln bezüglich der Helligkeits- und Sättigungsdifferenzen bei gesättigten Farben. *Farbe*, 18, 131-148
- Strocka, D. (1971), Color difference formulas and visual acceptability. *Appl. optics* 10, 1308-1313
- Wyszecki, G. (1968) Recent agreements reached by the colorimetry committee of the Commission Internationale de l'Eclairage. *J. Opt. Soc. Amer.*, 58, 290-292

DISCUSSION

Hunt: 1. Does your expression for ΔL correspond to the use of the Munsell value scale or the W^* scale for approximately evaluating differences between greys?

2. You refer to results from 2° -with dark surround, and 10° -with light surround conditions: might not the greatest tipping of ellipsoids occur with small fields in light surrounds? Have you any data for 2° -with-light-surround conditions?

MacAdam: 1. No, it does not. For different chromaticities of constant Y or W^* , L varies. For constant chromaticities, i.e. constant ratio of P/Q , e.g. a series of achromatic colors L is a linear function of $\log Y$, whereas W^* is a linear function of $Y^{1/3}$.

2. I do not have any data for ellipsoids for 2° fields in light surrounds, so I cannot judge. However, there was no essential difference between 2° fields in dark (Brown-MacAdam) and 10° fields in light surround (Brown-12-observers), so I am confident that 2° fields results for light surrounds would not show any significant tipping.

Wyszecki: Our recent determination of color-matching ellipsoids indicate a similar random distribution of tilt in the x, y, Y diagram. However, my interpretation of this randomness deviates from yours. I think that the observed (subthreshold) color-matching data are inherently too inconsistent to be useful

for testing the tilt. A tilt may still exist and may reveal itself by performing experiments on superthresholds.

MacAdam: This comment does not ask for an answer. However, I agree that the randomness of the tipping of the experimental ellipsoids is too great to justify a conclusion concerning tipping. However, if the ellipsoids should be tipped as much as indicated by the CIE-1964, the Munsell renotation or the cube-root formulas, the experimental results should show a consistency of tipping of which there is no indication in the results of Brown and MacAdam, nor, according to Wyszecki, in the results he has obtained with Mr. Fielder. I have no doubt that superthreshold judgments of equality of large color differences will require tipping of ellipsoids, but I do not consider that relevant for evaluation of either perceptibility of small color differences or the specification of tolerances of color matching.

Friele: In commenting on Dr. Hunt's remark I would say that the change of the Weber type discrimination to de Vries-Rose might result in a tilt for dark colours.

In taking the FMC-1 formula for optimizing several sets of data on colour discrimination, I would suggest to use it in a form in which the original parameter f is incorporated. f came out to be zero in the Chickering optimization, and is therefore omitted in the simplified form as presented by MacAdam. However the parameter f is wanted to account for possible rotation of chromaticity ellipses as compared to the original MacAdam ellipses.

MacAdam: I agree, and will reexamine Friele's parameter f in any future reoptimizations.

Reilly: Is it possible that any indication of tipping in your ellipsoids would be predominately in the same direction as the major axis of the ellipsoids and due to the derivation of the ellipses?

MacAdam: Anything is possible, but I see no indications of this in the available results. The directions of the major axes of the ellipsoids are not invariant with transformation of the chromaticity diagram, whereas the direction of displacement of the highest point on the ellipsoid is. See the discussion of this subject in Brown, MacAdam (1949).

K. Richter: This morning I referred to experiments of Dr. Kowaliski. His experiments show that we need a higher luminance factor for colours in the yellow region than for the achromatic surround to get colours of equal lightness. This is shown in your Fig. 4.

MacAdam: I am aware of the discrepancy between Dr. Kowaliski's results and the indications of Fig. 4, and cannot explain them. However, experimental results reported by Dressler, MacAdam, Wyszecki, Sanders and Breneman, and many other have shown that little if any greater luminance of yellow is required than of white to match the brightnesses, whereas significantly less luminance is required of other high-purity color, and extremely less of violet and violet purple.

K. Richter: You, Dr. Persels, Dr. Stiles and others used a line element. You get a mathematical expression for describing different colour attributes, such as colours of equal hue and equal brightness. Do you see any chance to get a mathematical expression to describe other colour attributes by the same line element, such as colors of equal blackness which are different from colours of equal brightness and colours of unique hues?

MacAdam: In addition to formulas for color difference (what you call a line element) Persels and the others who have derived loci of constant hue and brightness needed and used a *concept*, that of Schrödinger, that those loci were identical with geodesics. We have no similar concepts for blackness or unique hues. Until we have, no one can compute those attributes from fundamental data, or from color-difference formulae.

Stiles: In the line element formula given on p. 164 and p. 165 do the expressions satisfy the Schrödinger algebraic criterion for obtaining consistent equibrightness surfaces?

MacAdam: I do not know. The criterion specified by Schrödinger is very complicated and I do not know how to test the FMC-1 formula analytically for conformity to it. However, a numerical test indicates that the geodesic between two non-complementary spectral colors having the values of Y indicated in Fig. 4 attains the same values of Y at its intersections with the several loci in Fig. 4 as the values of Y found for those chromaticities in the original determinations of those loci. As I understand Schrödinger's discussion, this is what his condition implies and requires for the validity of the surface of constant brightness.

Malkin: 1. Were the minor axes of the Brown-MacAdam ellipsoids not tilted more than the major axes?

2. What is the effect of illumination level on the size and tilt of the discrimination ellipsoids obtained?

MacAdam: 1. My answer is essentially given to Reilly. I did not notice any greater tilt of the minor than the major axes. In any case, they are critically dependent on the particular representation (the $x, y, 0.2 \log_{10} Y$ space) and would be entirely different in any other of the infinite variety of permissible representations. They therefore have no inherent significations. Only the displacement of the chromaticity of the point of highest Y from the chromaticity of the center of the ellipsoid is invariant with transformations of the chromaticity diagram. Only that merits consideration in discussions of tilting of the ellipsoids.

Billmeyer: I am concerned about the several variations now available for calculating FMC-1 or FMC-2 color differences — for example, your earlier simplifications of the polynomials for K_1 and K_2 , and the several modifications mentioned in oral report today.

We know that the ellipses of FMC-1 differ from those of PGN by factors of

14 DISCUSSION

at least 2 in places, and FMC-2 differs by similar amounts from results from Simon Goodwin charts.

1. What magnitude of changes will the modifications mentioned today make (excluding of course effects of changing lightness weighting)?

2. For uniformity of practice in industrial use of these color difference formulae, what version would you recommend (excluding the FMC-1 vs FMC-2 controversy)?

MacAdam: 1. The formulas given in this paper are exactly equivalent to the two FMC-formulas, as discussed in the text.

2. I do not recommend any change of industrial use of the FMC-1, or FMC-2 formula (whatever is preferred by the user). I merely point out that both can be expressed much more simply than in their previously published forms, and I suggest that this simpler form of the FMC-1 formula be adopted by the CIE, after the 10 numerical parameters are reoptimized by use of whatever body of ellipsoid data is preferred by the CIE.

15 NEW COLOR-MATCHING ELLIPSES*

GUNTER WYSZECKI AND G. H. FIELDER

*National Research Council of Canada
Ottawa, Ontario, Canada*

Color-matching ellipses and color-matching ellipsoids of the kind first published by MacAdam (1942) and Brown and MacAdam (1949) have been used to develop and check empirical and inductive line elements of color space. (for a review, see Wyszecki and Stiles, 1967a) Some of these line elements are employed as a means of calculating color-differences in industrial color-control tasks.

In this paper we shall summarize the results of new experiments on color matching and their correlation with earlier data obtained by other investigators.

The salient details of our experiment are as follows: The 7-field colorimeter described previously (Wyszecki, 1965) was used with only two juxtaposed fields in operation. Each field is a regular hexagon which at the position of the observer, who uses both eyes in viewing, subtends approximately 3° . The two fields are viewed against a large (40°) white surround mounted in front of the instrument. The configuration of the visual field as it appears to the observer is indicated in Fig. 1.

Twenty-eight test colors were selected whose (x,y) -chromaticity points scatter over the chromaticity gamut provided by the instrument. Fig. 2 illustrates the gamut of the instrument and the distribution of the chromaticity points. Each test color was produced by an appropriate mixture of the three instrumental primaries serving the field on the left as seen by the observer. The luminance of each test color was set equal to $12 \text{ cd}\cdot\text{m}^{-2}$ whereas that of the surround was maintained at approximately $6 \text{ cd}\cdot\text{m}^{-2}$.

The spectral radiance distributions of the corresponding primaries of the two fields of the colorimeter are almost identical. Thus for each test color,

* This lecture is based on a recently published paper in J. Opt. Soc. Amer. 61, 1135-1152 (1971).

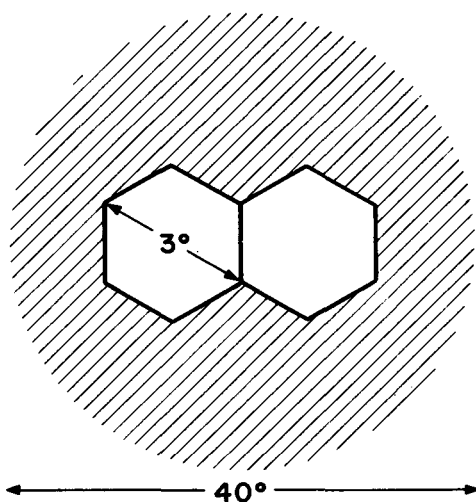


Fig. 1. Configuration of visual field. Hatched area is the white surround (not to scale).

presented in the left field, there can be obtained a matching color in the left field, whose spectral radiance distribution is almost identical to that of the test color. This ideal match is practically non-metameric and thus invariant to differences between color-matching functions of different observers.

Three observers (GF, AR, GW of ages 32, 27, 42 respectively) with normal color vision and extensive experience in visual colorimetry took part in the experiment.

A test color was set in the left field (Fig. 1) and the observer was asked to produce a matching color in the right field by manipulating the controls of the three primaries for that field. The instrumental readings were automatically recorded whenever the observer indicated that he had obtained a match. He then destroyed his settings and repeated the task. Typically a session was completed after the observer had made 30 matches on the given test color. Usually two or three sessions a day involving different test colors were scheduled for each observer.

In making his observations the observer located his forehead against a head-rest but his eyes were allowed to scan freely across the visual field; no arrangements were made to have strict fixation on a point or the fine dividing line between the two fields.

The instrumental readings obtained by an observer for his repeated matches of a given test color were converted into tristimulus values by taking into account the calibration data for the instrumental scales Wyszecki (1965). These values were in turn transformed into CIE chromaticity coordinates x, y and a luminance-related coordinate $l = 0.2 \log_{10} Y$. An appropriate statistical analysis (Wyszecki and Stiles, 1967b) leads to a color-matching ellipsoid in (x, y, l) -space.

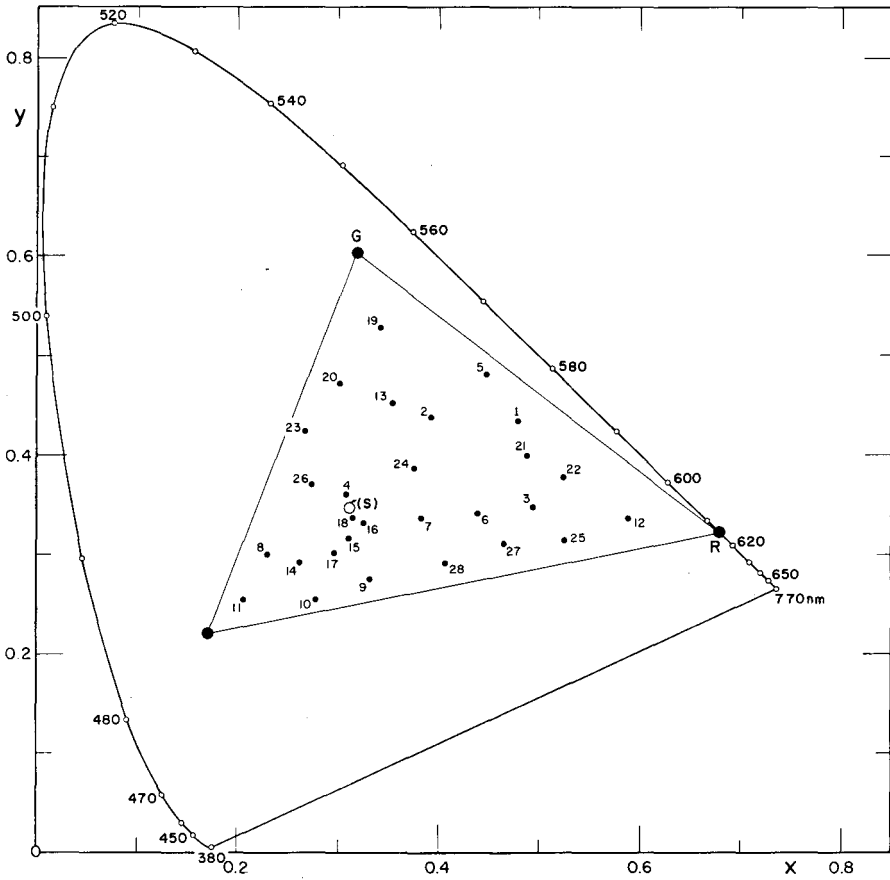


Fig. 2. CIE 1931 chromaticity diagram showing chromaticity points of the primaries (R), (G), (B) of the colorimeter, the surround (S), and the 28 test colors. The triangle (R), (G), (B) represents the chromaticity gamut provided by the colorimeter.

The equation of a color-matching ellipsoid is of the following form:

$$\begin{aligned}
 &g_{11} (x-x_0)^2 + g_{22} (y-y_0)^2 + g_{33} (l-l_0)^2 \\
 &+ 2g_{12} (x-x_0) (y-y_0) + 2g_{13} (x-x_0) (l-l_0) \\
 &+ 2g_{23} (y-y_0) (l-l_0) = (ds)^2
 \end{aligned}$$

From the g -coefficients, resulting from the statistical analysis of the observational points, and the constant $(ds)^2$ we can determine the size, shape, and orientation of the ellipsoid whose center is placed at the point (x_0, y_0, l_0) representing the test color.

Fig. 3 illustrates in (x, y, l) -space, as an example, the distribution of the 30 colors each of which GW observed as a "match" with a given test color during a particular session. Also shown are the three main projections and one cross-

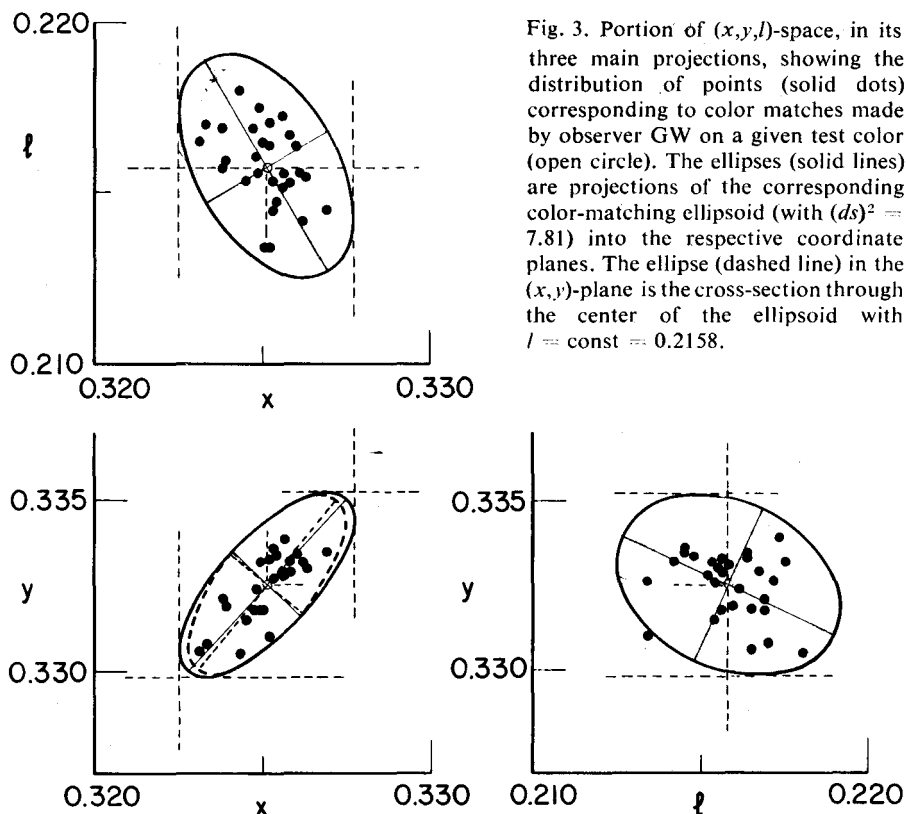


Fig. 3. Portion of (x, y, l) -space, in its three main projections, showing the distribution of points (solid dots) corresponding to color matches made by observer GW on a given test color (open circle). The ellipses (solid lines) are projections of the corresponding color-matching ellipsoid (with $(ds)^2 = 7.81$) into the respective coordinate planes. The ellipse (dashed line) in the (x, y) -plane is the cross-section through the center of the ellipsoid with $l = \text{const} = 0.2158$.

section of the color-matching ellipsoid derived from the observed color matches.

The constant $(ds)^2$ is set equal to 7.81. This value corresponds to the χ^2 -value for 3 degrees of freedom and stipulates that the ellipsoid contains, on the average, 95% of any random set of color matches.

The coefficients g_{11} , g_{12} , g_{22} , and $(ds)^2$ of each ellipsoid determine an ellipse which is the cross-section, $l = \text{const}$, through the center of the ellipsoid. This ellipse can be used as a measure of the observer's precision of matching the chromaticity of the test color. The observer's precision of chromaticity matching is considered the more important aspect of the color-matching experiment and we will confine our discussion to the elliptical cross-sections of our color-matching ellipsoids and will call them "color-matching ellipses" or just ellipses. These can be compared readily with the results of similar investigations by other authors.

Fig. 4 shows the elliptical cross-sections derived from the color-matching ellipsoids of observer GW. Similar results were obtained for observers GF and AR. The orientation, shape, and size of the ellipses vary with the location of the center of the ellipse in the chromaticity diagram largely in accordance with

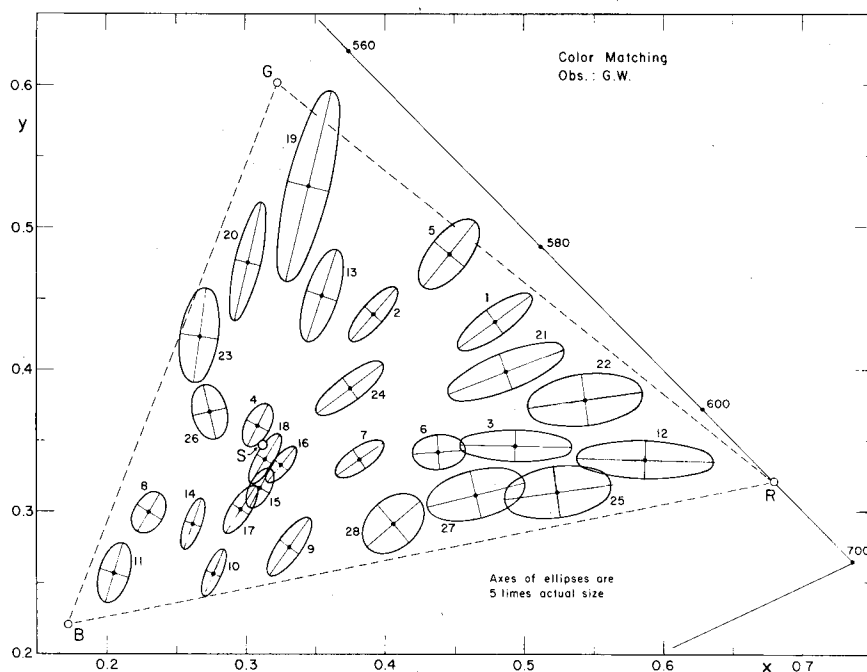


Fig. 4. Portion of CIE 1931 chromaticity diagram showing cross-sections ($l = \text{const } 0.2158$) of color-matching ellipsoids obtained in (x, y, l) -space for observer GW. Points R, G, B represent the chromaticities of the instrumental primaries and S the surround.

expectations based on the experimental results obtained by MacAdam (1942), Brown and MacAdam (1949), Brown (1957) and Stiles (1946)*. However, in some instances significant deviations from one set of ellipses to another are indicated.

The intercomparison of the different sets of ellipses of earlier publications with those obtained by us is somewhat restricted because of the smaller chromaticity gamut we had available. Nevertheless, important trends of agreement and disagreement can be found within the common gamut in the central area of the chromaticity diagram.

There is an overall resemblance between the different sets of ellipses. On the other hand, for a given color center the ellipses of different observers are not always in close agreement. Occasionally rather large discrepancies are noted in orientation, size, and shape. No doubt, some of these discrepancies arise from the inconsistencies inherent in the ellipses of the individual observer as

* The ellipses of Stiles are predicted by his 'modified Helmholtz line element' and not directly observed. However, for the purpose of describing our intercomparison of sets of ellipses of different origins, we will refer to Stiles' ellipses as "WSS ellipses" and to Stiles as "observer WSS".

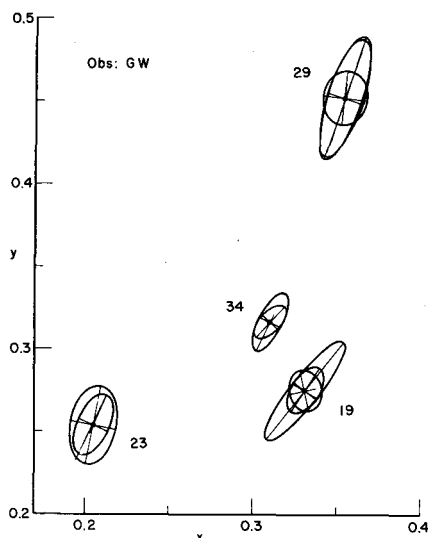


Fig. 5. Portion of CIE 1931 chromaticity diagram showing cross-sections ($l = \text{const} = 0.2158$) of color-matching ellipsoids obtained in (x, y, l) -space for the same observer (GW) producing sets of color matches at different occasions but otherwise identical observing conditions.

exemplified in Fig. 5. Other discrepancies must be attributed to systematic differences between the visual mechanisms of different observers and to differences in the observing conditions used by different investigators.

The variability of color-matching ellipses based on repeated sets of observations, as shown in Fig. 5, is remarkable and deserves special attention. The statistically estimated standard deviation of the major and minor axes (a and b) of a color-matching ellipse, determined from 30 matches, is only ± 13 per cent. Visual color-matching data obtained by the same observer on different occasions seem to be considerably less repeatable than statistics predict. It appears that the functioning of the visual mechanism is affected by parameters that depend upon time and possibly other events and circumstances not accounted for in the model underlying statistical inferences.

Agreement or disagreement between corresponding ellipses obtained by different observers is readily displayed in the correlation diagrams shown in Figs 6 to 9. These diagrams correlate the orientation, the logarithm of the size, and the shape of corresponding ellipses. The orientation is defined as the angle θ of inclination of the major axis (a) of an ellipse relative to the x -axis of the chromaticity diagram. The size is given by $A = \pi ab$, and the shape is indicated simply by the ratio a/b of the major and minor axes.

The agreement between the orientation (θ) of the corresponding ellipses of a given pair of observers shown in Figs 6 and 7 is generally quite good.

The agreement between the size (A) of the corresponding ellipses of a given pair of observers is also fairly good. However, all of the ellipses of WSS (Stiles, 1946) are considerably larger than the corresponding ellipses of PGN (MacAdam, 1942). On the average they are larger by a factor of $(4.5)^2 = 20.25$

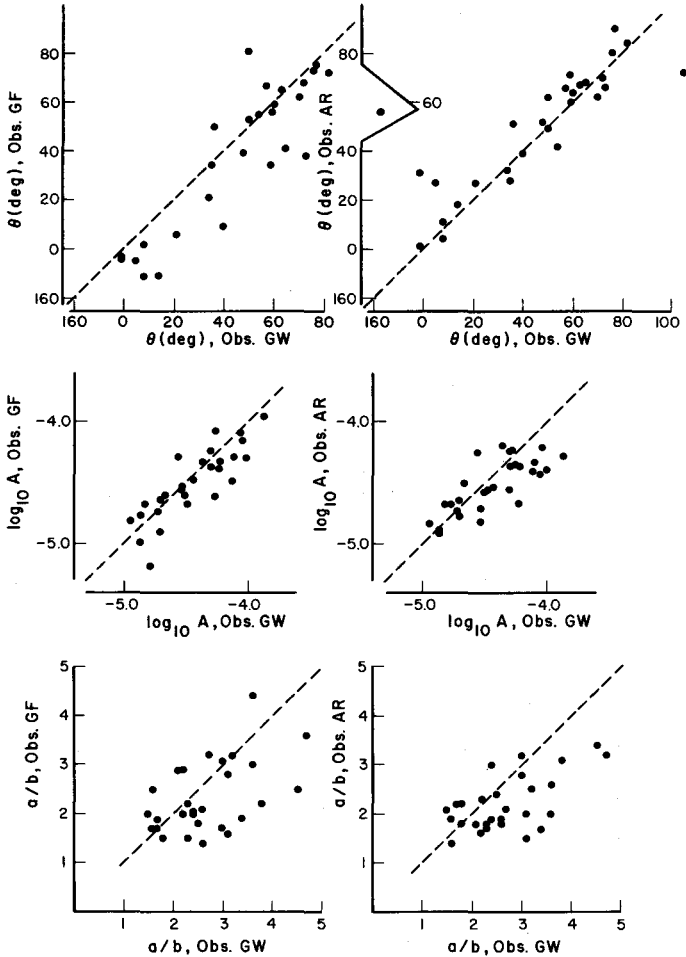


Fig. 6. Diagrams showing the correlation of the orientation (θ), size ($\log_{10} A$), and shape (a/b) between corresponding color-matching ellipses of observers GF, GW, and observers AR, GW.

which has been indicated in the $\log_{10} A$ -diagram by a dotted line displaced from the 45° -dashed line by $\log 20.25 = 1.31$.

The agreement between the shape (a/b) of the corresponding ellipses of any pair of observers is poor and few if any inferences regarding correlation can be made. The generally poor correlation between the shape of corresponding ellipses reflects to a large extent the inherent inconsistencies in the data of individual observers. These inconsistencies are amplified by the fact that we compare the ratios of the semi-axes a and b each one of which being subject to considerable uncertainty (see Fig. 5).

Correlation diagrams of the kind shown in Figs 6 and 7 cannot be drawn

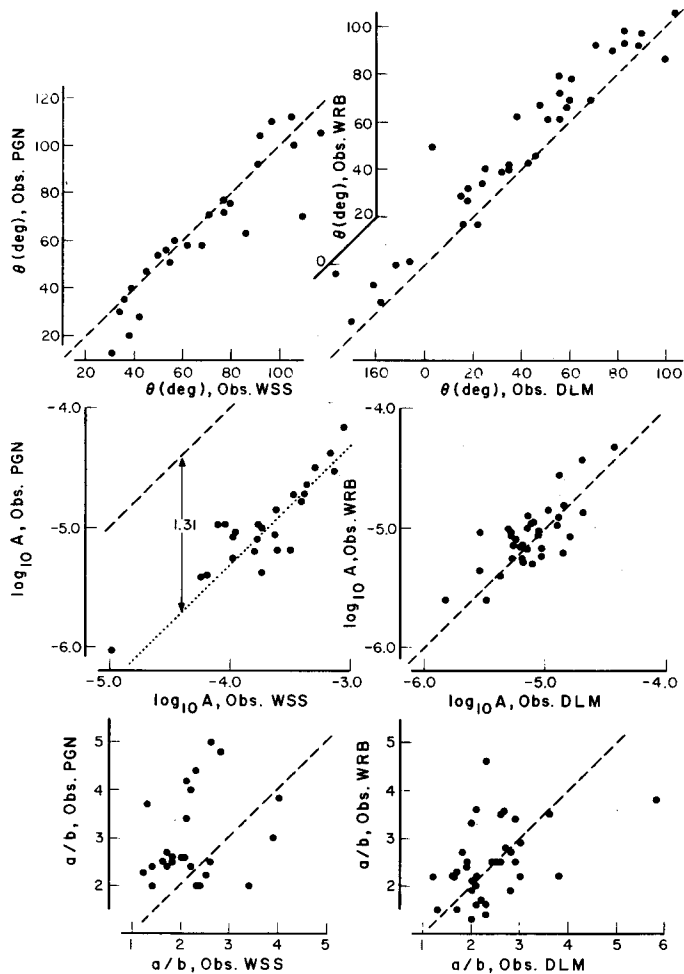


Fig. 7. Diagrams showing the correlation of the orientation (θ), size ($\log_{10} A$), and shape (a/b) between corresponding color-matching ellipses of observers PGN (MacAdam, 1942), WSS (Stiles, 1946) and observers WRB, DLM (Brown and MacAdam, 1949)

readily for observers whose ellipses do not directly correspond to one another, that is do not refer to the same test colors. However, in making use of the Friele-MacAdam-Chickering formula (Chickering, 1967), we can establish a link between such pairs of observers. The parameters of the FMC formula were optimized such that computed FMC ellipses would agree as closely as possible with the corresponding observed ellipses of PGN. Fig. 8 (left-hand side) illustrates the agreement between the orientation, size, and shape of the two sets of ellipses. The agreement is quite good for the orientation (θ) and size ($\log_{10} A$),

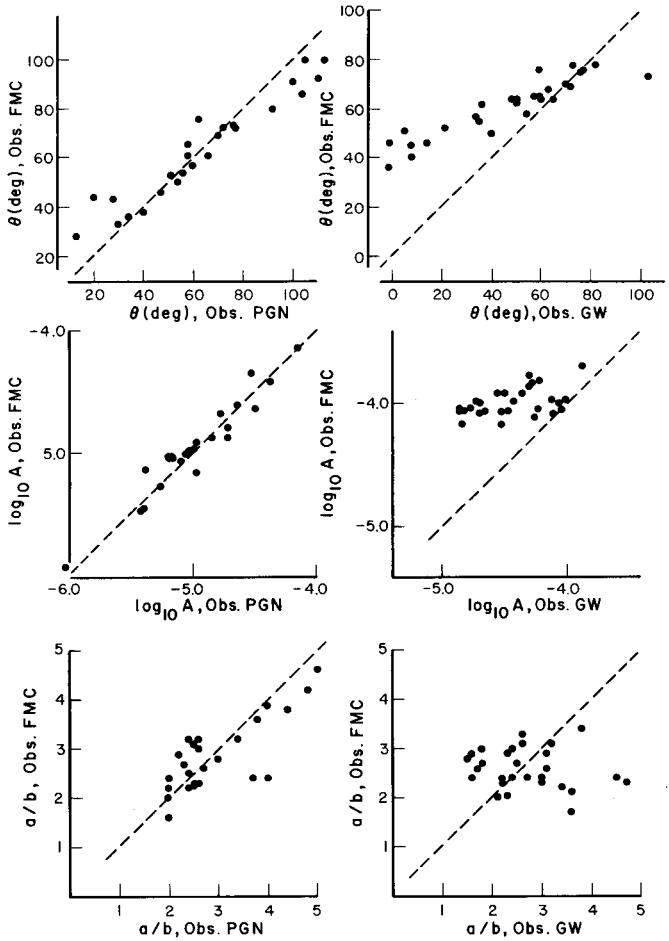


Fig. 8. Diagrams showing the correlation of the orientation (θ), size ($\log_{10} A$), and shape (a/b) between corresponding color-matching ellipses of observers FMC (Chickering, 1967), PGN (MacAdam, 1942), and observers FMC, GW.

not so good for the shape (a/b). We accept the FMC ellipses as representative of the PGN ellipses. Furthermore we postulate that FMC ellipses computed for other color centers will also be representative of observations that could have been made by observer PGN.

Fig. 8 (right-hand side) shows correlation diagrams for observers FMC*

* We realize that FMC is not an "observer" in the true sense of the word, but for reasons of simplicity we will use it in describing our intercomparison of sets of ellipses of different origins.

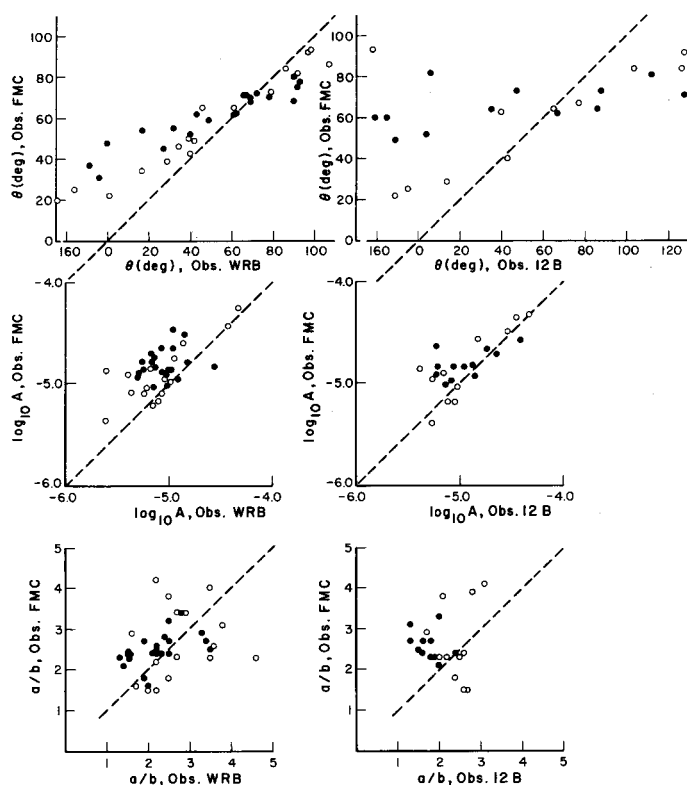


Fig. 9. Diagrams showing the correlation of the orientation (θ), size ($\log_{10} A$), and shape (a/b) between corresponding color-matching ellipses of observers FMC (Chickering, 1967), WRB (Brown and MacAdam, 1949), and observers FMC, 12B (Brown, 1957).

and GW, and Fig. 9 similar diagrams for FMC and WRB, and FMC and 12B. These diagrams reveal important systematic discrepancies between the ellipses of FMC and those of GW, WRB (Brown and MacAdam, 1949), and 12B (Brown, 1957; mean of 12 observers). The most apparent discrepancy occurs in the orientation θ . In the θ -correlation diagrams the points do not distribute along the dashed 45°-line but rather along a less inclined line. In taking recourse to the actual ellipses we find that the systematic discrepancy in orientation applies to the purple and red test colors. In this area of the chromaticity diagram the FMC ellipses (representing observer PGN) are systematically more steeply sloped than those of observers GW, WRB, and 12B. Since observer GW correlates well with GF and AR, and observer WRB correlates well with DLM, we conclude that PGN shows similar discrepancies also with observers GF, AR, and DLM. The only observer that correlates well with PGN is WSS.

Systematic discrepancies between FMC and GW are also indicated with re-

gard to the size of corresponding ellipses. Whereas the GW ellipses range in size over a full log unit (approximately), those of FMC range over 0.4 log units (approximately). The discrepancies between FMC and WRB, and FMC and 12B are not as pronounced.

We also notice differences between the average size of the ellipses of one observer as compared to that of the ellipses of another observer. In particular we recall the large discrepancy between PGN and WSS shown in Fig. 7. However, the factor of $(4.5)^2 = 20.25$ (that is, $\log_{10} A = 1.31$) can be explained, at least in part, by the fact that PGN's ellipses refer to "standard deviations" of color matches which MacAdam (1942) estimates to be roughly one-third of the corresponding "just-noticeable" color differences. On the other hand, the WSS ellipses are "threshold" loci based on a 50% chance-of-seeing a color difference. In addition, PGN viewed steadily a 2° bipartite field in a white surround, while WSS viewed very short flashes (0.063 sec) of a 1° test field superimposed on a 10° field with a dark surround.

A lesser but still significant discrepancy is shown between the average size of the GW ellipses and that of the FMC ellipses (Fig. 8). Better agreement on the average would be obtained if the size of each FMC ellipse would be reduced by 0.4 log units, that is, by a factor of 2.5 (approx.). We recall that the FMC ellipses represent observer PGN who viewed a 2° field in a white surround, while observer GW viewed a $3^\circ \times 6^\circ$ field also in a white surround. The increased field size for observer GW would tend to increase his precision of color matching and thus reduce the sizes of his ellipses relative to those of PGN.

The average sizes of the corresponding ellipses for observers FMC and WRB (Fig. 9) show a relatively small discrepancy of about 0.2 log units, that is a factor of 1.6 (approx.). Those of FMC are larger by that factor. The FMC ellipses refer to 2° visual field-white surround conditions, those of WRB to 2° visual field-dark surround conditions. Whether the differences in surround can explain the small discrepancy in average size of ellipses is difficult to say. A genuine observer difference of that magnitude is also possible.

In the case of observers FMC and 12B, again the average size of the FMC ellipses is somewhat larger than that of the 12B ellipses. However, the discrepancy is rather small and only 0.1 log units, that is a factor of about 1.25. Actually, we would have expected a larger discrepancy here because the 12B ellipses refer to 10° observations as compared to 2° observations for the FMC ellipses.

It is also possible to intercompare the average sizes of the ellipses of, for example, observers GW and WRB (Figs 8 and 9). The average size in the case of GW is $\log_{10} A_{av} = -4.4$ (approx.), that of WRB is $\log_{10} A_{av} = -5.0$ (consider only solid dots of Fig. 9). The discrepancy is 0.6 log units, that is a factor of about 4. We recall that the ellipses of GW were calculated with $(ds)^2 = 7.81$, while those of WRB with $(ds)^2 = 1$. Thus, if everything else would be equal,

we should expect a discrepancy by a factor of 8 (approx.) rather than 4. However, taking into consideration the difference in visual field used by the two observers ($3^\circ \times 6^\circ$ in white surround for GW, and 2° in dark surround for WRB), we are inclined to reduce the average size of the GW ellipses by a factor of about 2. Indeed this would reduce the discrepancy from a factor 8 to a factor 4.

The distribution of points in the a/b -correlation diagrams is again rather unfavorable and do not permit us to draw conclusions as to any agreement or lack of agreement between different sets of ellipses.

The systematic discrepancies between the ellipses of FMC and those of every other observer (with the exception of WSS) are similar in nature, particularly with regard to the orientation θ . This similarity leads us to conclude that the ellipses of observers GF, AR, GW, WRB, DLM, and 12B all correlate quite well in orientation, but those of PGN and WSS, while correlating well with each other, appear to stand isolated from the rest.

The sets of ellipses of observers WRB, DLM, and 12B cover a larger chromaticity gamut than those of GF, AR, and GW. In intercomparing these sets we considered only those ellipses that fall into the gamut illustrated in Fig. 2. In the correlation diagrams of Fig. 9 the solid dots refer to ellipses inside this gamut, the open circles refer to those outside. However, the distinction drawn here has only little, if any, effect on our conclusions.

We cannot offer any quantitative explanation for this split of the observers into two distinct groups other than to suggest that different experimental procedures may have contributed to it significantly.

We recall that observers GF, AR, GW, WRB, DLM, and 12B all operated three control knobs of a three-primary colorimeter and determined color-matching ellipsoids from a trivariate distribution of color matches. On the other hand, observer PGN made color matches by turning a single control knob so as to vary the color of one half of the visual field along a straight line in the chromaticity diagram, automatically keeping the luminance constant. The standard deviations of color matches along a series of straight lines intersecting at the chromaticity of the test color results in the color-matching ellipse whose center coincides with the chromaticity of the test color. WSS used his two-color threshold technique to derive the parameter functions of his line element from which, in turn, color-matching ellipses can be calculated for any color center.

The question arises whether the use of three fixed primary colors in producing color matches introduces inadvertently a bias in the distribution of the color matches for a given test color. We notice the tendency of the color-matching ellipse to orient itself toward the chromaticity point of the "nearest" primary. The striking similarity of orientation of the new ellipses presented here and those obtained by Brown-MacAdam (1949) and Brown (1957) may be due to the

fact that the different experiments involve primaries of similar chromaticities. On the other hand, the experiments made by observers PGN and WSS do not involve mixtures of three primaries fixed throughout the experiments. Obviously, further tests are required to resolve the discrepancy between the two groups of data.

Perhaps the main points to be made on the basis of our study are the following:

- i) Visual color-matching data obtained by the same observer on different occasions, but identical observing conditions, appear considerably less repeatable than is indicated by normal statistical analysis. The functioning of the visual mechanism seems affected by parameters that depend upon time and possibly other events and circumstances not controlled by normal statistical methods.
- ii) In allowing for inconsistencies in the color-matching ellipses of individual observers, we find that our new sets of ellipses for observers GF, AR, GW correlate quite well with those for observers WRB, DLM, and 12B reported by Brown and MacAdam in 1949 and Brown in 1957.
- iii) The ellipses (particularly for red and purple test colors) of observer PGN (MacAdam, 1942) show significant and systematic discrepancies from those of observers GF, AR, GW, WRB, DLM, and 12B. The discrepancies are puzzling and further experiments appear necessary to explain them. We suggest that differences in experimental procedure may, in part, be responsible for the discrepancies.
- iv) The ellipses predicted by the FMC formula agree quite well with those of observer PGN. This was to be expected as the parameters of the FMC formula are optimized parameters to obtain specifically the best fit with the PGN data. What is surprising is the fact that the ellipses predicted by Stiles' line element (observer WSS) agree with those of observer PGN almost as well as those of FMC with PGN. This agreement must be viewed also by keeping in mind that no numerical optimization of the parameters of Stiles' line element was made to obtain specifically a best fit with the PGN data.

We now recall that Stiles' line element has been constructed on fundamentally different arguments than the FMC formula, with respect to the visual mechanism that governs color discrimination (Wyszecki and Stiles, 1967b). This then leads us to conclude that good agreement between a set of ellipses predicted by a given line element (such as FMC or Stiles) and a set of corresponding color-matching ellipses (such as those of PGN) obtained experimentally is a necessary, but obviously not sufficient condition to make the given line element acceptable from the point of view of vision theory. Additional tests on the line element, involving experimental data other than

color-matching ellipses, are required to establish the acceptability of the line element.

- v) The different observing conditions used in the different experiments investigated here are expected to have an effect on the orientation, size and shape of the color-matching ellipses. However, the only effect we were able to extract with some certainty from the discrepancies between different sets of ellipses, was the change of the absolute size of the ellipses for different visual field sizes. Color-matching ellipses increase in size with a decrease in visual field size.

The differences in luminance of the test colors and the luminance and chromaticity of the surround in the different experiments were not of sufficient magnitude to reveal a significant effect they may possibly have on the orientation, size, and shape of the ellipses. Further experiments, specifically designed to study the effects of these variations in observing conditions, are required.

REFERENCES

- Brown, W. R. J., and MacAdam, D. L. (1949), Visual sensitivities to combined chromaticity and luminance differences. *J. Opt. Soc. Amer.* 39, 808-834
- Brown, W. R. J. (1957), Color discrimination of twelve observers. *J. Opt. Soc. Amer.* 47, 137-143
- Chickering, K. D. (1967), Optimization of the MacAdam-modified 1965 Friele color-difference formula. *J. Opt. Soc. Amer.* 57, 537-541
- MacAdam, D. L. (1942), Visual sensitivities to color differences in daylight. *J. Opt. Soc. Amer.* 32, 247-274
- Stiles, W. S. (1946), A modified Helmholtz line element in brightness-colour space. *Proc. Phys. Soc. (London)* 58, 41
- Wyszecki, G. (1965), Matching color differences. *J. Opt. Soc. Amer.* 55, 1319-1324
- Wyszecki, G., and Stiles, W. S. (1967), Color science. Concepts and methods, quantitative data and formulas. New York, etc., Wiley, 1967
- a) pp. 511-560
- b) pp. 536-541

DISCUSSION

Hunt: When the lengths of the major and minor axes of the ellipses are not very different, the angle of the ellipse is not of much significance. What effect does this have in the comparisons of the angles of pairs of systems of ellipses?

Wyszecki: It is true that for near-circular ellipses the orientation of the major axis is of little or no significance. However, in the x,y chromaticity diagram this rarely occurs.

Hunt: Does the variation in angle of the ellipses show more about the nature of x,y chromaticity diagrams than about the differences between the systems of ellipses?

Wyszecki: I don't think so. When transforming the x,y diagram to a uni-

form chromaticity type diagram in which the ellipses would tend to become circular, the correlation of the orientation of corresponding ellipses may appear less pronounced but the inherent similarity that was noted earlier would not be lost by the transformation.

Bartleson: Your 40° surround had a luminance of about 6 cd.m^{-2} and the test fields were maintained at about 12 cd.m^{-2} . I believe this level is approximately 1/2000 of that encountered in daylight. Would you care to comment on the advisability of extrapolating these results to higher luminances and to other test-to-surround luminance and spatial relationships?

Wyszecki: Indeed, as compared to daylight luminance levels our luminance levels were very low and strictly speaking an extrapolation from our low levels to daylight levels cannot be made. Experimental data for high levels are not available and this should be one of the next steps to be taken in future experiments to expand our data.

Nimeroff: 1. Should you not have tried to apply the variances and covariances for the spectral tristimulus values of the standard observer to see if the resultant ellipses correlate with those for the selected colors?

2. The selected color technique requires interpolation to obtain variability data for intermediate chromaticity coordinates.

Wyszecki: Test and comparison colors in our field of view are nearly non-metameric and the use of the CIE standard observer instead of the color-matching functions of the actual observer is justified. Nevertheless it would be of interest to analyse our data also in terms of the variability of the spectral tristimulus values (color matching functions) of the individual observer.

Walraven: From the short glance on the graph you showed of your own new results, it seems that the crispening effect is larger than in the Nutting data. Is that the main reason for a low correlation with the Nutting data?

Wyszecki: No, the crispening effect does not appear to have upset the correlation. Instead, the main discrepancies between corresponding ellipses are in the purple and red region.

Guth: In regard to your earlier comment that some kinds of averaging causes ellipses to be circles, how do you know that in fact the circles are not the "true" contours and that samples from the true case yield artifactual ellipses?

Wyszecki: I do not know which one of a set of repeated ellipses for the same test color is the "true" ellipse or whether the "true" ellipse is more circular or less circular than an actually observed ellipse.

Clarke: Many people in industry working on practical problems question whether colour discrimination or various data measured with a visual colorimeter will really apply to surface colours, as it is difficult to make the mode of appearance look, not like an "aperture colour", but convincingly like a surface colour. What are people's opinions on this?

Wyszecki: If you adjust the luminances of the surround and test color to be

about equal or the surround luminance to be higher than that of the test color very often the surface-color mode is observed. However, I agree that actual field trials with real object colors are always recommended to verify the visual colorimetric data.

Terstiege: You did your color-matching experiments with white surrounding fields, that means daylight or illuminant C. Therefore you got smaller ellipses at the white C chromaticity point than in the more reddish region of the color diagram. What would happen to your experiments by using a surrounding field of illuminant A? Or even a red surround?

Wyszecki: By changing the color of the surround, we anticipate a decrease in the size of the ellipses where centers are near or at the chromaticity of the surround. This phenomenon is related to the "crispning" effect mentioned earlier.

Boynton: In making a color match, there are many options available to the subject concerning his approach to the task. Each subject probably has his own style, although the approach of a given subject probably also varies from time to time. Was the subject constrained in any way with respect to the manner in which he converged upon a color match?

Wyszecki: We did not restrain the observer in any way with respect to the measures in which he obtained his match, except to ask him to destroy completely the match before proceeding to repeat the match. Also we did not detect any systematic pattern of approach the observer used to arrive at a match when starting from a given color far removed from the matching color.

MacAdam: I wish to commend Dr. Wyszecki and Fielder for producing this fine body of data. It is a significant and badly needed contribution to solution of the problem of establishing a method for specifying the perceptibility of small color differences. We need more of such work and data.

By way of comment on the several questions, objections and suggestions for modification or extensions of Wyszecki and Fielder's work, I express the opinion that anyone who had done such an enormous job must have already thought of all those matters and must have good reasons for taking the alternatives that they had. It is therefore unreasonable to expect Wyszecki and Fielder to do their work over again with the suggested modifications. Rather, let those who think they could do this work better, undertake and do it for themselves. We need and would welcome all additional, comparably reliable, results.

Simon: 1. Did you throw out any experimental data?

2. Were the ellipsoids tilted systematically towards the neutral axis?

Wyszecki: 1. A few were rejected which were quite obvious erratic and presumably caused by temporary instrumental failure of recording the setting.

2. There were tilts but with no apparent systematic orientation. I believe the data are not suitable to extract proof or disproof of tilt.

Indow: In your diagram showing fluctuation of ellipsoids of an individual

observer, all the ellipsoids coincide in the center of gravity. Is it because you disregarded the fluctuation of means in plotting?

Wyszecki: This is correct. The statistically expected deviations between the different means did occur, but we made the means to coincide exactly with the given test color.

16 CONTINUATION OF THE STUDY OF COLOUR THRESHOLDS

F. PARRA

*Laboratoire de Physique, Muséum National d'Histoire Naturelle,
Paris*

In Stockholm I presented data (Parra 1969) on the whimsical shape of colour difference threshold contours in the chromaticity diagram. These data are presented in Fig. 1. Most data were already reported five years ago (Parra 1966), but the typical characteristics are confirmed by later experiments, e.g. in the red corner of the diagram (Parra 1969).

The threshold contours are obtained by measurements of the colour difference thresholds in thirteen directions. The luminance level was about 300 cd/m^2 (1200 trol.) and the field of view $3^\circ 50'$. These was either a dark or a light surround, the latter of a luminance of about 100 cd/m^2 (400 trol.). The measurement were completed by one observer. New data were obtained three years later and these showed essentially the same features. These features are that the contours of the colour difference threshold in the chromaticity diagram are not smooth ellipses, but that they have extensions in the directions of the confusion points of protanopes and tritanopes. The contour in the blue corner of the chromaticity diagram shows these features most characteristically.

When the saturation discrimination data of Priest and Brickwedde (1936) are reconstructed in the form of a discrimination contour (Fig. 2), the results are comparable with ours. Other preliminary data show that some observers have extensions in the direction of the deuteranopic confusion point.

During the discussion in Stockholm, doubts were raised about the significance of the star-shaped extensions. In particular Dr. Vos asked whether I could make a simple and therefore convincing demonstration of my thesis that "just noticeable areas" can be concave. Such a concavity means that in Fig. 3 a mixture of two colours *A* and *B*, both not discriminable from *O*, can give a mixture *C* which is visibly different from *O*. This question can be answered now.

The first remark to this is however the asymmetry in the directions *OA* and

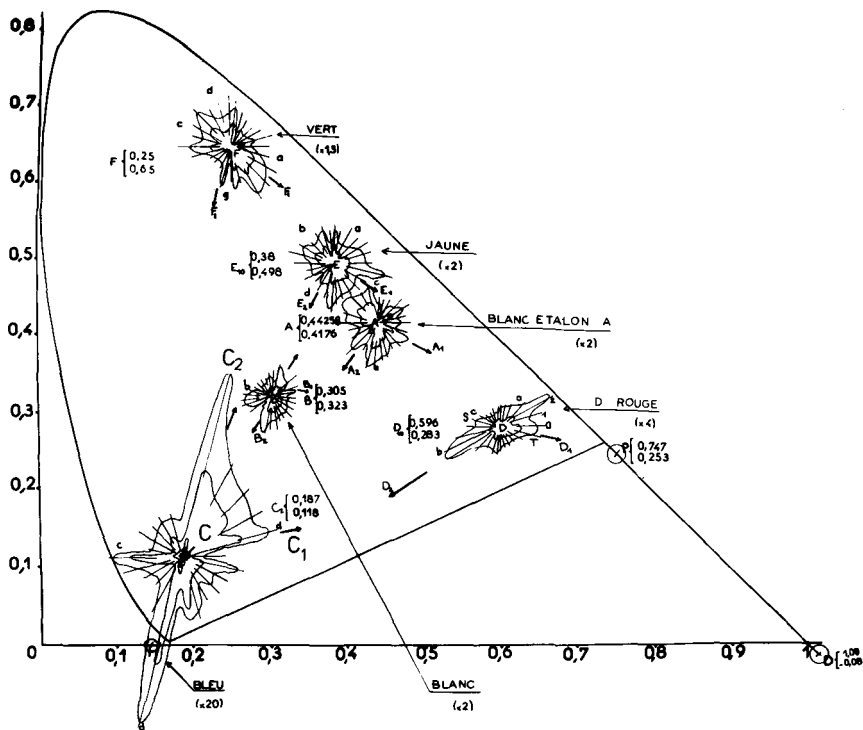


Fig. 1. Colour difference threshold contours around six colour centres. The size as drawn is a factor 20 times the actual size. Attention is called for the directions CC_1 and CC_2 indicated by arrows in the "blue" contour, towards the protanopic confusion centre P and the tritanopic confusion centre T . The outer contour is obtained by augmenting, the inner contour by diminishing the JND contour with the standard error σ .

OC , the difference in length on both sides of O . This point has been discussed extensively in relation with the data of Crawford, also presented in Stockholm (Crawford 1969). This is not yet explained. Some students of mine repeated experiments of this type and found the same differences.

The problem of the curvature of the threshold contour is even more difficult. Dr. Clarke mentioned during the discussions in Stockholm how tedious a good calibration of a colorimeter is. Therefore especially the setting of the sectors in my apparatus have been checked and some deficiencies of the colorimeter are corrected. This modification has been carried out only recently and the modified colorimeter has not yet been extensively employed for actual measurements. I agree with Dr. Vos that theoretically point C should not be different from the colour in point O , but not enough measurements have yet been done to confirm this.

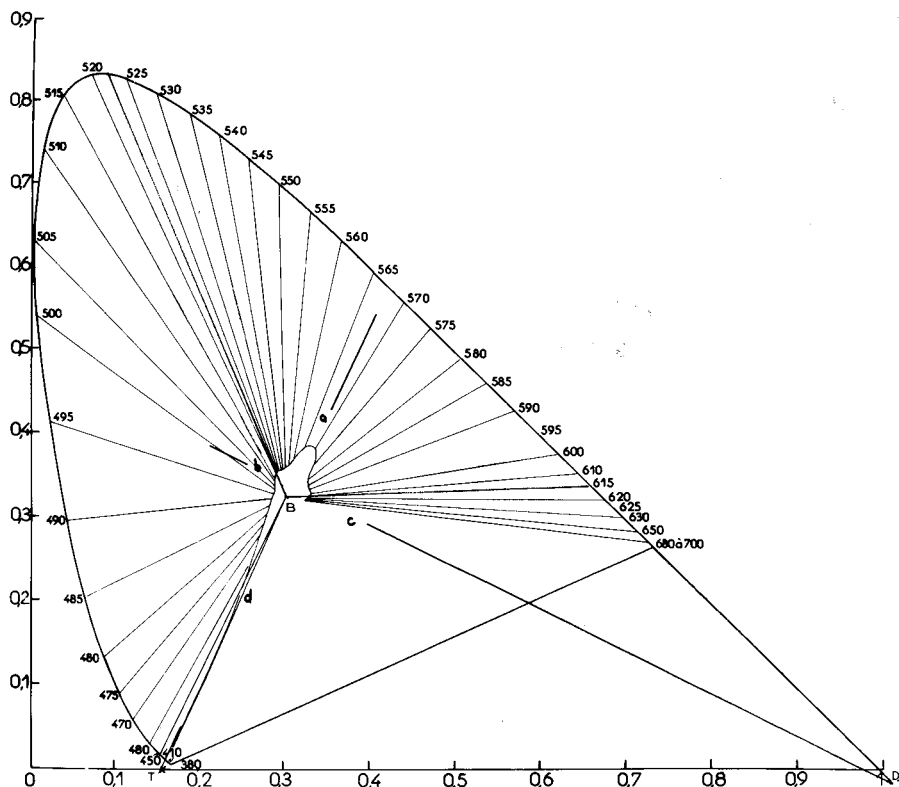


Fig. 2. Purity difference thresholds around the white point B according to Priest and Brickwedde (1938). The contour is obtained by transformation of their results to x, y data. Note the concavity of the contour.

REFERENCES

- Crawford, B. C. (1969), Just perceptible chromaticity shifts. Proc. 1st AIC congress "Color 69", Stockholm, 302-311
- Parra, F., (1966) Recherches sur le seuil différentiel de couleur. Thesis University of Paris
- Parra, F., (1969), Seuils différentiels de couleur. Proc. AIC Stockholm, 727-733
- Priest, I. G. and Brickwedde, F. G. (1938) The minimum perceptible colorimetric purity as a function of dominant wavelength. J. Opt. Soc. Amer. 28, 133-139

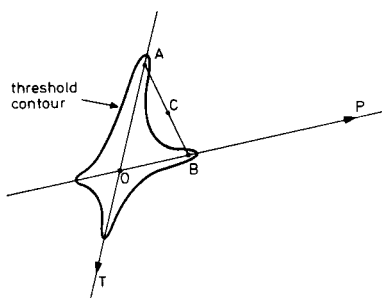


Fig. 3. The concave shape of the threshold contour predicts that mixture of A and B, both non-discriminable from O, can produce a colour C outside the contour, and thus noticeably different from O. Preliminary experiments indicate that C is indeed different from O, but the significance of these findings is not yet good enough to be sure.

DISCUSSION

Vos: I do not know whether I serve the dubious merit to be taken for the only opponent in Stockholm. I think we were with many, and in particular Ruddock had some doubts about your apparatus. Please will you, in the proceedings, describe the apparatus and the experimental procedure in detail. New and unconventional results simply require extra care in argumentation.

Parra: I chose to refer to your question in particular since your question seemed to be most pertinent. I did not give details on the equipment as I presumed them to be known. I will be happy to furnish the information as soon as I will be able to present new and unquestionable data, obtained with my improved equipment.

Stiles: Did you use frequency of seeing curves to determine your thresholds and if so, what was the shape or steepness of these curves? The shape and steepness of probability of seeing curves is a factor in determining whether threshold measurements will lead to a sum of squares line element—like $ds^2 = \Sigma dx^2$ —and to discrimination ellipses. If the probability of seeing curves are very steep, however, an element with sums of higher powers would be expected— $ds^n = \Sigma dx^n$ —and the ellipses would fill out almost to rectangles for large n .

Perhaps your results would suggest that, on the contrary, a value of n smaller than 2, even smaller than 1 is operative, connected perhaps with very shallow frequency of seeing curves of low precision of threshold determination, or with some kind of super-summation.

Parra: No, I did not use the frequency of seeing method. I just changed the setting in certain direction, until the difference with the comparison field became just visible. The indicated disparity is the root mean square value of the disparities found. Your proposal suggests an interesting experimental method, though, which should be tried.

Jaekel: A very simple thought has occurred to me. I wonder whether it is relevant to explain the apparent contradiction between conventional ellipses and Dr. Parra's findings. Dr. Wyszecki commented over differences between the Nutting and other ellipses and wondered whether methodology was involved, i.e. how matches are obtained. He also mentioned that observers got tired and said "this is a good enough match". Does Dr. Parra not ask a different question?

When two stimuli are mixed in a particular ratio, do they look different from another stimulus? Can the ellipse assessors be asked if their appropriate extreme matches are different from the test stimulus?

Will they still say "This is a match?" If not, the different findings are due to different psychological initial frames of reference.

Parra: One cannot exclude fatigue, of course, but why in preferential directions in particular?

MacAdam: I understood Dr. Jaekel to ask (or to suggest) whether the task

of the observer was different (and easier) when he reported that the difference *OC* in Dr. Parra's Fig. 1 was perceptible than when he reported that the differences *OA* and *OB* were not perceptible. On the basis of Dr. Parra's statements that *A* and *B* were obtained by adjustments made by the observer, whereas *C* was produced by the experimenter (by additive combination of *A* and *B*) I suggest that the observer's task was easier in the case of *C* and that he was more likely to report that he perceived the difference *OC* (even though it was less perceptible than *OA* and *OB*) because he would not, in the case of *C* be called upon to readjust it, as he would be if he reported either *OA* or *OB* perceptible.

Parra: I don't think the question comes up with this type of experiment, as *C*, obtained as it is by mixing *A* and *B*, cannot in one trial be made to appear different or non different from *O*, as could be done with *A* and *B*. I hope the definitive paper will show this.

Clarke: Dr. Parra mentioned that the other observers in his experiment were dichromats. Are *you* what is known as a "normal trichromat"?

Parra: I showed up normal on Farnsworth 100 Hue test. Should I doubt its validity? The curves shown are for normal vision. It is hardly necessary any more to demonstrate that dichromatic curves stretch in the direction of the confusion axes. Nevertheless I checked some color blinds and they did.

To me the starshape in the two confusion directions indicates that normals are to be considered as double-dichromats indeed.

Verschuere: I cannot understand how the midpoint *C*, lying between *A* and *B* which are both inside the 1 MacAdam unit ellipse, can fall outside the ellipse.

Parra: But that is just the point. I don't believe the JND contour to be an ellipse.

Nimeroff: It may be, because the question asked the observers, that their data determine the adjacent ellipses about the test color. The ellipse is a convenient representation of the scatter of data points, because it has two axes and an orientation and is a closed figure. The ellipse is usually imposed on data; we should rather let the data speak for themselves.

Kowaliski: Two of Prof. Parra's students who confirmed his observations during their postgraduate work have also acted as observers in our determinations of equivalent luminances. To this purpose their normality of vision was carefully checked and found to be perfect, as well as their reliability as observers. This adds confirmation to Prof. Parra's results.

Wagner: The present discussion has been about the rather extraordinary shape of Dr. Parra's threshold "ellipses". It seems to me, that their contours fit nicely into Dr. Guth's concept of superadditivity. His assumption of inhibition of opponent color-mechanisms might apply to the present data as well.

I. NIMEROFF

*National Bureau of Standards
Washington, D.C., U.S.A.*

Investigators and users of color space metrics always feel the need to search for a color metric that will yield a "uniform" color space. The space currently in vogue to satisfy this need is the 1964 CIE $U^*V^*W^*$ system. This space does, however, have limitations. Perhaps the most serious restriction is that the spectrum locus in this system shifts with the luminance level and with the chromaticity coordinates of the adapting source, thereby resulting in a multitude of spectrum loci, perhaps an infinite number of loci. This paper, with its somewhat frivolous title, is a plea that when a new color space is being considered for international recommendation that it be thoroughly and critically examined for possible limitations to its use.

In 1931 the CIE adopted a Standard Observer and Coordinate System for Colorimetry (See Judd, 1943) which was recommended for subsequent use in expressing color data so that results would be readily comparable. Since that time this color system has been very widely used, with considerable success, in color matching, color prediction, and color and color tolerance specification. The success of this system results from its simplicity in being an all positive number system, rather than a positive and negative number system, such as the WDW system (See Wright, 1946). The 1931 CIE Standard Observer and Coordinate System has clearly defined color regions, as has been pointed out by Kelly (1943). It has a monotonic Planckian radiator locus and a well defined, orderly spectrum locus. All of these characteristics can be represented conveniently on a plane.

In recent years colorimetrists have been seeking a metric in which perceptually equal color difference magnitudes will be represented equally in a "uniform color space." Their original ideas of what constitutes a uniform color space could not be made to conform to the experimental facts of perceptually uniform color differences (See Judd, 1970). The reason for this difficulty is that the mag-

nitude of the color difference between two colors is influenced by the color of the background on which the colors are observed. The experiments indicated that a uniform color space can be found only for colors viewed with a background that differs in color minimally from the colors that are being compared. Application of such a requirement, however, would lead to the maintenance of a huge library of colored backgrounds on which to make color difference judgments for all colors in practical use. Because extension of the original ideas of uniform color space was required, a decision was reached to make such judgments of color differences on a neutral background of approximately the same luminance as the colors that are being judged. Luminance has been found to be the most important of the color parameters on which a visual color-difference judgment is based. Hence, a color space in which luminance of test colors and background has a dominant role is a requisite.

Through the years there have been developed a number of suggestions for uniform-space color systems. Most of these have been established to make uniform some special region or regions of the color space, while other regions have been neglected or left distorted in the process. Generally, chromaticity diagrams of these spaces do not differ either practically or theoretically from each other. Schultze (1969) has been studying the effectiveness of some eight, at first, and later twelve, of the color difference formulas on which these color difference spaces are based. Each of these spaces was found to have its characteristic difficulties, particularly for highly saturated red and blue colors that differ in lightness and saturation. Schultze has recently found that the uniform color space, currently recommended by the CIE, the $U^*V^*W^*$ system, "gave the worst agreement by far." This poor agreement may well be a consequence of characteristics of this system that are discussed in this paper.

The CIE $U^*V^*W^*$ space.

This space, proposed to the CIE by Wyszecki (1963) and recommended by the CIE in 1964, has a value function based on a study by Ladd and Pinney (1955). The equations for this space are:

$$\begin{aligned} U^* &= 13W^* (u - u_0) \\ V^* &= 13W^* (v - v_0) \\ W^* &= 25Y^{1/3} - 17, \end{aligned} \tag{1}$$

where u , v and u_0 , v_0 are respectively, the chromaticity coordinates of the test color and the adapting source color in the 1959 CIE u , v , w uniform color space and Y is the luminance of the test color. Thus, for each of the many possible chromaticities of adapting sources there is a $U^*V^*W^*$ space.

In a recent study conducted at the National Bureau of Standards by Haupt and Schleter (not yet published) the loci of chromaticity coordinates of combinations of Lovibond glasses were developed by computation for sources A

and C. The results of this study were similar to those found by Werner (1968) for Corning glass filters.

Spectrum loci of the $U^*V^*W^*$ space

The problem with the $U^*V^*W^*$ system is that there are no rules given about how to treat the value function W^* in determining the location of the spectrum locus. There are several options available for such rules by which W^* can be defined, with or without any restrictions. We have considered two of these options. First, we let Y , the luminance, take on constant values for all wavelengths. In the first option, we let Y take on several constant values between 0.0 and 100.0. In the second option, Y was made proportional to the spectral luminous efficiency function $V(\lambda)$. It should be obvious that with each of these options different spectrum loci will be developed.

Before considering the effects of these options in deriving the spectrum loci let us reexamine the basis of this $U^*V^*W^*$ system. It is primarily the CIE 1959 u, v, w , uniform color spacing system. If we consider the chromaticity coordinates of spectrum colors $u(\lambda), v(\lambda)$ and of the adapting source u_0, v_0 (See Nimeroff, 1964), we can plot their differences to obtain the spectrum locus. We see that we have plotted the u, v chromaticity diagram with its origin translated from 0,0 to u_0, v_0 . In other words, Eqs (1) have become:

$$\begin{aligned} u(\lambda) - u_0 &= U^*/13W^* \\ v(\lambda) - v_0 &= V^*/13W^* \end{aligned} \quad (2)$$

and we have obtained the U^*, V^* chromaticity diagram for $13W^*$ equal to unity, or $Y=0.319$. Thus, for various values of Y we may obtain various values of W^* or $13W^*$ to obtain for each a different spectrum locus. Several values of Y, W^* , and $13W^*$ are tabulated below to illustrate how they are related.

Y^*	W^*	$13W^*$
0.0	-17.0	-221.00
0.314	0.00	0.00
0.319	+0.077	+1.00
1.00	8.00	104.00
10.0	36.85	479.05
100.0	99.05	1,287.65

By using these values of $13W^*$, we may obtain several spectrum loci, as is shown in Fig. 1.

For the second option, we chose $Y(\lambda)$ proportional to the spectral luminous efficiency function $V(\lambda)$, thus, $Y(\lambda) = KV(\lambda)$. The spectrum loci have been developed for values of K equal to 1, 10, and 100. These have been shown in Fig. 2. With this option it can be seen that the spectrum loci for K equal to 1 and to 10 have overlapping values and the space is not clearly defined by a plane. The number of loci when Y is a function of the spectral luminous efficiency function

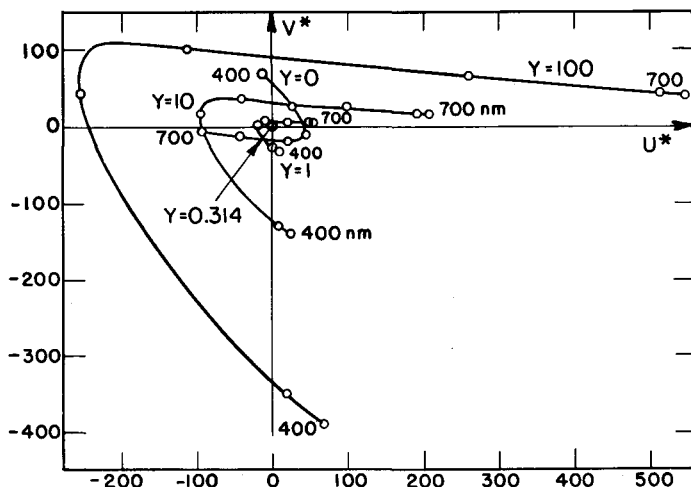


Fig. 1. CIE spectrum loci of U^* , V^* for $Y = 0.0, 0.314, 1.0, 10.0$, and 100.0 .

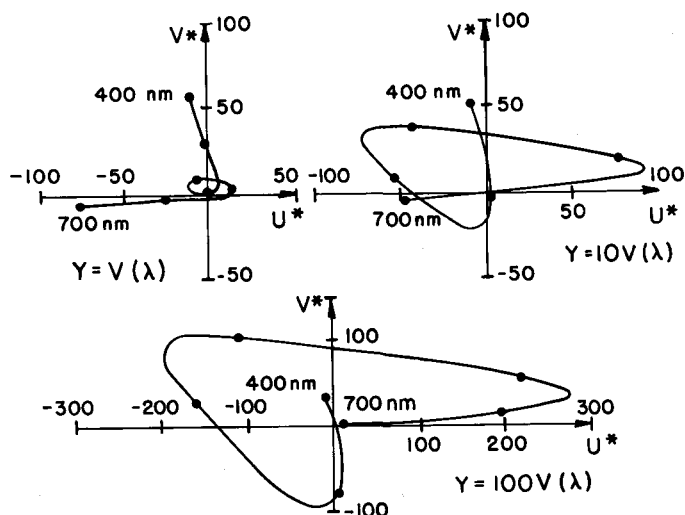


Fig. 2. CIE spectrum loci of U^* , V^* for $Y(\lambda) = V(\lambda), 10V(\lambda)$, and $100V(\lambda)$.

is determined by the number of values of K that are used. A plot of W^* against U^* reveals that there are very large changes in U^* with relatively small changes in W^* , particularly for saturated red and blue colors.

On the basis of this discussion we see that even with a sensible selection of rules there is no one spectrum locus for the $U^*V^*W^*$ system but an infinite number of loci. We should, however, attempt to restrict ourselves to a limited number of loci by a proper restriction of useful rules.

Statistical considerations.

In developing a useful color and color difference metric we must keep in mind a fundamental principle in measurement science that is based on statistical principles. We make measurements in order to find a likely value. This process requires that we take sufficient data not only to establish that value but to attach an uncertainty to it. To obtain such an uncertainty we must have at least one more measurement than there are parameters in the metric. In other words, the data should be sufficient to have at least one redundancy. This means that in a system defined by an equation in which there are N parameters and N or less data are obtained the uncertainty of the metric in that system would be infinitely large. We would then be obliged to consider the metric to be meaningless.

Recommendations.

It is evident that the shortness of the interval that elapsed between the proposal of the $U^*V^*W^*$ system in 1963 and its adoption in 1964 provided insufficient time for its critical review. Adoption to a recommended procedure should be made only after thorough and critical examination to determine its capabilities and limitations to its use. We close this paper with a plea that any metric for color difference and color measurement that is being considered for international recommendation be kept to a reasonably simple form, so that excessive amounts of data would not be needed to obtain meaningful results.

REFERENCES

- Judd, D. B. (1933), The 1931 I.C.I. Standard Observer and Coordinate System for Colorimetry. *J. Opt. Soc. Amer.* 23, 359-374
- Judd, D. B. (1970), Ideal Color Space. *Color Eng.* 8 (2) 37-52
- Kelly, K. L. (1943), Color Designations for Lights. *J. Opt. Soc. Amer.* 33, 627-632
- Ladd, J. H., and Pinney, J. E. (1955), Empirical Relationships with the Munsell Value Scale. *Proc. IRE.* 43, 1137
- Nimeroff, I. (1964), Spectral Tristimulus Values for the CIE (u, v, w) Uniform Spacing System. *J. Opt. Soc. Amer.* 54, 1365-1367
- Schultze, W. (1969), Umfassender Vergleich Von Sieben Farbabstandsformeln. *Proc. 1st AIC Congress "Color 69"*, Stockholm, 621-640; *Farbe* 18, 105-130
- Werner, A. J. (1968), Luminous Transmittance and Chromaticity of Colored Filter Glass in CIE 1964 Uniform Color Space. *Appl. Optics* 7, 849
- Wright, W. D. (1946), *Researches on Normal and Defective Colour Vision*. Kimpton, London
- Wyszecki, G. (1963), Proposal for a New Color-Difference Formula. *J. Opt. Soc. Amer.* 53, 1318-1319

DISCUSSION

Wyszecki: I would advise that we adhere to the recommendations in detail and recall that these include a number of restrictions to the $U^*V^*W^*$ system which, it appears, you have not fully expressed in your paper.

Nimeroff: I have read the recommendations but others have not. $U^*V^*W^*$ was devised for reflecting materials. When one starts to apply this color difference system to other uses, as is being considered in a CIE committee for signal lights, namely to use $U^*V^*W^*$ to represent color differences among signal lights, we must reexamine the system. This paper, therefore, should not be considered as a criticism of the CIE $U^*V^*W^*$ system of evaluating color difference. It is, rather, meant to be a caution not to use the system for such problems for which it was not intended.

Hunter: I want to plead for fitting color-measurement systems specifically to different modes of appearance. The CIE Y, x, y color measurement system fits the aperture mode. The $U^*V^*W^*$ system is intended for surface colors. For transparent films and liquids, and for bare metals, scales which expand more in chromaticity at low values of Y are needed.

I was with Dr. Judd and Mr. Priest when they came back from the 1931 CIE meeting and I remember that we said: we want something we can apply to materials. They said: we have to start at the beginning, specifying the most basic things first, i.e. the color of an aperture—and the CIE system does it beautifully. We need however to set up some ground rules for going beyond apertures.

Billmeyer: It is with regret that we learn from Dr. Nimeroff of the recent death of Francis C. Breckenridge. For many years he was one of the very few proponents of the Breckenridge-Schaub x'', y'' transformation as a uniform chromaticity diagram for signal light purposes. This transformation is of course, well behaved with respect to spectrum locus and, being an attempt to circularize the MacAdam-Nutting ellipses is appropriate for use with aperture colors. It was used for many years by CIE E-1.3.3., colors of signal lights, but has now been replaced (perhaps inappropriately) by the U, V diagram and by extension the $U^*V^*W^*$ system. Clearly, a uniform chromaticity (trichromatic) system suitable for use with aperture colors (signal lights, transparencies, etc.) is needed, as Mr. Hunter has pointed out in this discussion. What if any, are the current suggestions or recommendations for such a system?

Nimeroff: I agree, that there is need for such a system, but there are currently no other suggestions. One additional comment. In 1968 we published a monograph, which was really an updating of Judd's circular on colorimetry, in which we deprecated the discussion of color difference. People criticized us for that—this showed us that people had read the document!—and noticed that we had omitted most of that discussion.

The reason was that color difference equations were then—and still are in 1971—in such a state of flux that anything we would have said would be misleading and inconclusive really. Without knowledge of variability of a system we have nothing at all. With it we make the metrology a color measurement science. We really need simple equations—and $U^*V^*W^*$ are such simple

equations. If we formulate the proper restrictions on the system—and make sure these instructions are being read- and understood, we will have a very useful system for evaluating color difference.

Reilly: We have also occasionally encountered the same problem and have devised a simple expression for reflectances below 1% in which lightness is proportional to Y .

Nimeroff: Where there is a need for treating data that existing techniques cannot do, we must attempt to find a solution. It may be that a linear relation between W^* and Y , ($W^* = k Y$, $0 < Y < 1$) might give adequate weighting of the chromaticity differences, Δu , and Δv . Before recommending the use of such a relationship, however, it should be thoroughly checked.

Hunt: There really is a need for extending the use of $U^*V^*W^*$ below $Y=1$, not for surface colours, but for filter colours as in colour transparencies.

At the ISCC Williamsburg conference John Yule came up with a W^* extension below $Y=1$.

Apparently we cannot stop people to try to use it below $Y=1$, so we have to make some recommendations, I think.

18 EQUIVALENT LUMINANCES AND THE REPRODUCTION OF COLORS

P. KOWALISKI

*Kodak-Pathé
Vincennes, France*

One of the important parameters in color reproduction by photography, motion pictures and television is the third attribute of color, its brightness or lightness. The creation of an image involves indeed the reproduction of the entire tone value scale extending from black to white, with only a smaller part of the total information contained in the reproduction determined by the chromaticities, i.e. by the two other attributes of color, hue and saturation. The quality of color reproductions is therefore frequently evaluated mainly by the fidelity or adequacy of the reproduction of the luminance scales of the original, currently called "tone reproduction", and tone reproduction characteristics, while of course including color scales, are normally referred to the neutral or achromatic scales. Investigation of the tone reproduction in color scales (Kowaliski, 1963), however demonstrates its close relationship to the inherent brightness of the various hues made apparent in color order systems; on the other hand the tone scales in color depend on the brightnesses of colors at varying luminance levels, i.e. on their equivalent luminances.

THE EQUIVALENT LUMINANCES OF COLORS

In the colorimetric system most adequate for the evaluation of tone reproduction in color scales, the 1964 U^* , V^* , W^* system of the CIE, the colorimetric luminance is determined by the third parameter W^* which in turn is derived from the Y value of the 1931 CIE system, based on the choice of the spectral luminosity function $V(\lambda)$ for $\bar{y}$ as one of the three tristimulus functions. Ex-

perience having however shown that the sensation of brightness or lightness of a colored area is not related in a simple manner to colorimetric luminance, various researches were carried out in the past for the determination for the various hues of the ratio of brightness B or lightness L to the colorimetric luminance Y .

According to a recommendation of the Colorimetry Committee E-1.3.1 of the CIE of 1965, brightness and lightness can now be expressed in measurable luminance units as the "equivalent luminance" of colors. This parameter still corresponds to the definition of lightness, given by Judd as the attribute of a color that allows classification of its perception by equivalence with the perception of an achromatic stimulus, taken in a series ranging from black to white (Judd, 1952). The determination of the equivalent luminances of colors therefore requires the choice of a neutral reference scale of luminances and the successive matching of the steps of the color scale by heterochromatic photometry.

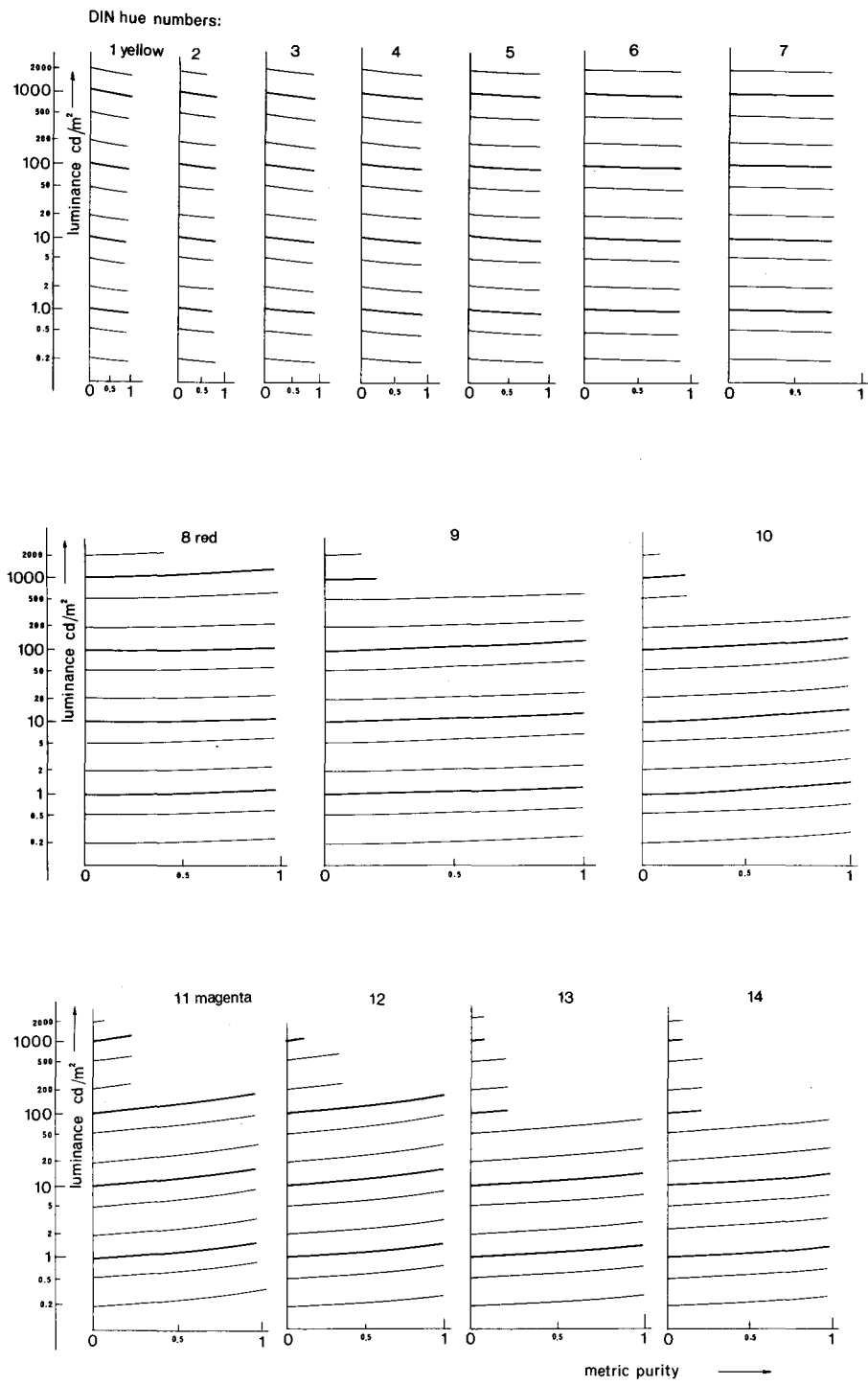
Results of determinations of the L/Y ratio were reported by Chapanis and Halsey (1955), Sanders and Wyszecki (1957, 1958) and Wyszecki (1967) and equivalent luminances were successively determined by Kowaliski first at rather low levels of luminance (1963), then up to 200 cd/m^2 (1967, 1969^a), and more recently, with a photocolormeter of very high luminance, up to $2,000 \text{ cd/m}^2$ (1969^b). These latter values cover the domain of the viewing conditions of color reproductions but still remain quite below the luminances of actual scenes. The open gate luminance of a motion picture screen, i.e. of its maximum white is 55 cd/m^2 , the minimum of the "white" of a television home receiver is 70 cd/m^2 and can reach values up to 140 cd/m^2 with "black surround" tubes (Benson, 1970), or even 170 cd/m^2 in color television home receivers (Hunt, 1967), and critical direct inspection of unprojected transparencies is standardized in several countries at $1100 \pm 300 \text{ cd/m}^2$ (USA, 1969). The luminance levels employed for the observation of photographic or paper prints can vary widely as shown by the figures given by Hunt (1965):

Approximate luminance (cd/m^2)

	White	Black
Sunlight in office	8,800	200
Daylight in office	340	10
Tungsten light in room	24	0.5

Here again standards specify rather high luminance levels such as $700 \pm 150 \text{ cd/m}^2$, which correspond to a rather well illuminated room, but insure critical evaluation of print quality.

The results of the present determinations of equivalent luminances, shown in the diagrams of Fig. 1, by giving results up to $2,000 \text{ cd/m}^2$, cover satisfactorily the viewing conditions of motion picture and television screens and also those of the inspection of paper prints under average conditions; they also



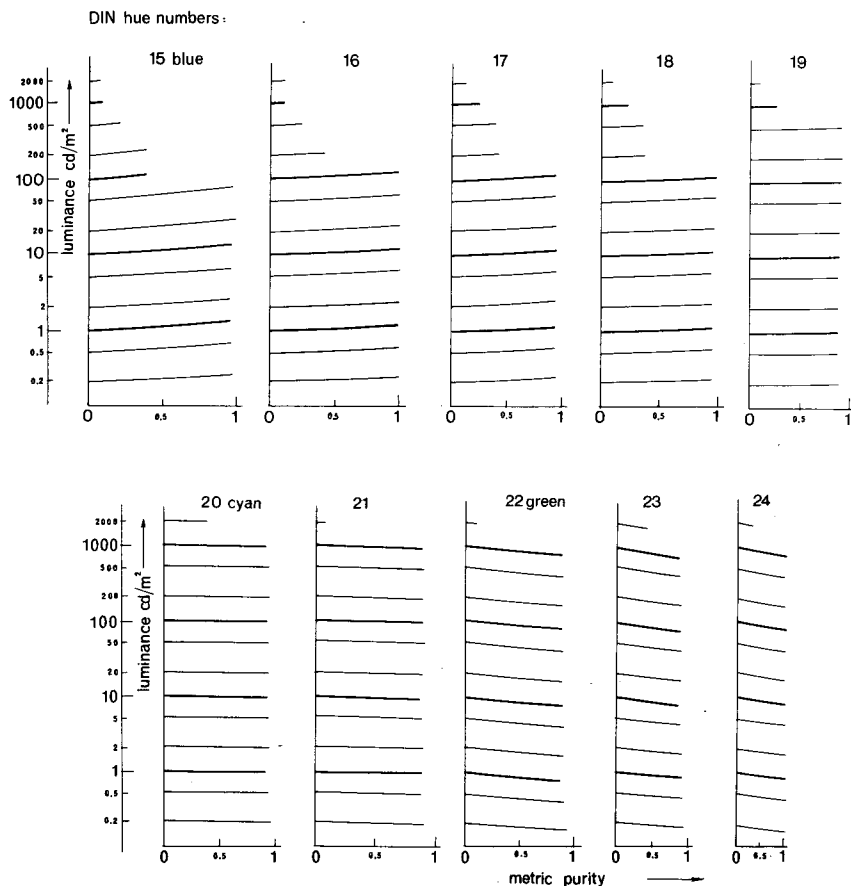


Fig. 1. Equivalent luminances for the colors of the twenty four DIN hues. These are luminance—purity plots showing for each hue and metric purity the average matching luminances of the neutral field.

apply to the standardized illumination levels for the viewing of color transparencies and prints, but as mentioned before still remain much below the luminance levels encountered during picture taking. For sunlit scenes these can be as high as $33,000 \text{ cd/m}^2$, and still reach several thousand candelas per square meter in professional photographic, motion picture, and television studios.

To attain luminance levels of this order in a photometer with a trichrome matching field of adequate size is not only difficult to achieve, but it appears not to be really useful as the levels attained in the present instrument already raise problems for the observers. With very high luminance levels of the matching field adaptation leads to a rapid saturation loss in the center of the

matching field, which besides this then also seems to fluctuate and to move, so that the central separation line becomes for some observers unstable and shows floating fringes in complementary colors. While reliable and reproducible results are however obtained by an adequate repetition of the matches by a sufficient number of observers, it does not seem useful to push the luminance levels still higher; the results of such measurements would besides this only extend the diagrams by about one log luminance unit and therefore only yield limited additional information.

The 24 diagrams of Fig. 1 are vertical planes of the generalized photocolorimetric grid previously described (Kowaliski, 1969^a) and shown in Fig. 2, similar to those proposed earlier by Brewer, Ladd and Pinney (1954). As all horizontal planes of the present cylindrical grid are u - v chromaticity diagrams, its radial vertical planes laid through the lines of specific chosen hues*) are metric purity—

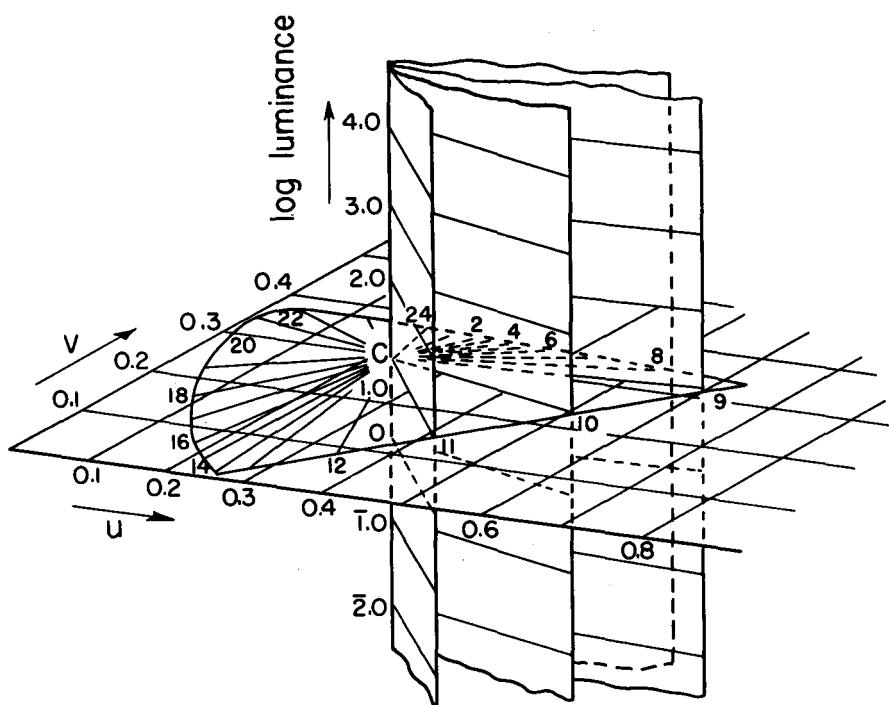


Fig. 2. Perspective view of the generalized photocolorimetric grid; all horizontal planes are 1960-CIE u - v chromaticity diagrams of identical size, so that the vertical planes through the central axis yield metric purity—luminance graphs.

* In all our determinations of equivalent luminances these hues were the twenty four hues of the DIN-system, according to a previous recommendation of the Committee "Bases théoriques de la Colorimétrie" of the French Association of Colorimetry.

log luminance plots. Positions of colors not lying on these planes, or of luminances intermediate to those chosen for the diagrams, can be determined by second order interpolation (Kowaliski, 1969^a); this interpolation was employed also for the computation of the curves shown in Fig. 1, corresponding to specific luminance levels. First the arithmetic means of the measured values are obtained, then the logarithms of these means, and the coefficients for the second order interpolation are then found by taking one higher (x_1) and two lower (x_2 and x_3) luminance levels of the same chromaticity u_i , v_i . Calling the means of the measured equivalent luminances y_1 , y_2 and y_3 , this yields

$$y_1 = x_1^2 a + x_1 b + c$$

$$y_2 = x_2^2 a + x_2 b + c$$

$$y_3 = x_3^2 a + x_3 b + c.$$

The interpolated values of the equivalent luminances are thus directly obtained in the computer (IBM 1130) and can be plotted in the grid.

Two important facts result from the inspection of the array of curves in the diagrams of Fig. 1:

First, the increase of equivalent luminance with saturation for most colors except the yellows, previously reported for lower levels of luminance from the other investigations, is again confirmed and extended to the higher luminance levels now reached.

Secondly, the ratio of equivalent luminance to nominal luminance in the achromatic scale remains constant under current conditions of daylight adaptation and therefore independent of the general luminance level. This second fact is of great importance for the reproduction processes because it demonstrates the possibility of observer adaptation to the satisfactory viewing of images, at comparatively low illuminance levels, of highly illuminated scenes and vice-versa.

THE EVALUATION OF COLOR REPRODUCTION CHARACTERISTICS

Most photographic or electronic reproduction processes serve pictorial representations intended to be judged by visual observation; among the well defined parameters customarily employed for the critical evaluation of the reproduction characteristics tone reproduction is therefore, as mentioned above, among the most important. Observer appreciation of the luminances and metric purities of the colors in the reproduction is accordingly a fundamental criterion of reproduction quality and the equivalent luminance diagrams of Fig. 1, containing information on the actual visual effects of color stimuli, thus prove very useful for quantitative evaluations of this kind,

As an example let us consider the photocolorimetric limits of reproduction processes as investigated by Brewer, Ladd and Pinney (1954) for color film

used as program material in television. In this investigation the limits were determined for a television screen and a color film in the three metric purity-luminance planes of the six hue pairs important in color reproduction processes, i.e. blue-yellow, green-magenta and red-cyan. The spectrophotometric characteristics of both the television screen phosphors and the color film dyes have changed since that communication was made and are now much improved. While thus not applying to present conditions, specially by the exaggeration of the differences between film and television, the Brewer, Ladd and Pinney data were here retained only as an example for demonstrating the usefulness of equivalent luminance information. The boundaries of their domains were therefore replotted in the corresponding planes of the equivalent luminance grid and are shown in Fig. 3 by the dashed curves.

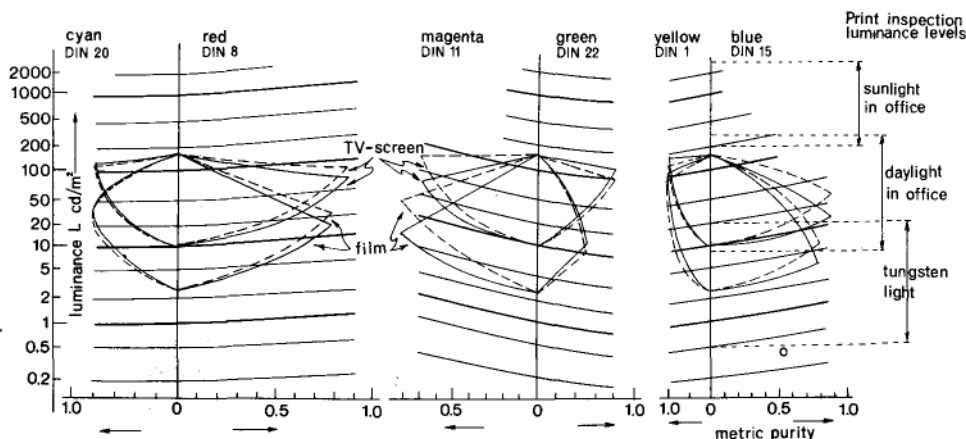


Fig. 3. Metric purity—log luminance diagrams showing the domains available for reproduction by television and by film. Dashed contours: uncorrected limits. Full contours: limits corrected for equivalent luminance. On the right hand the luminance levels for print inspection are also indicated.

These luminance and purity limits are not only determined by the spectrophotometric characteristics of the screen phosphors and film dyes, but they also depend to a great extent on the actual viewing conditions. As shown by Bartleson and Breneman (1967) the actual television image provides relative brightness reproduction close to the theoretically desirable 1 : 1 ratio, intermediate between that of a projection screen image and that of a reflection print. When the luminance scale of the film is adjusted for dark field surround and that of the television screen for dim surround, the two scales show actually much smaller differences than those apparent in Fig. 3.

In the computations by Brewer, Ladd and Pinney the available luminance range of the television screen was taken as limited by surface reflections and

roomlight to the value of 1.2 in the logarithmic neutral luminance scale, while the corresponding figure for the color film, in spite of the density range of its dye images of about 3.0, was 1.85 (The density and log luminance scales are of opposite sign but otherwise identical in size). This latter loss of luminance resulted from the flare present during projection and unwanted reflections in the projector. On the other hand the luminances of the saturated colors on the television screen were under these circumstances much higher than those reproduced by the color film; the film dyes, due to the only partial correction of their unwanted absorptions, were indeed darker than desirable for satisfactory luminance reproduction.

In the films employed at present these corrections are more complete, mainly by the so-called interimage effects; development reaction products, diffusing from one layer into the adjacent ones inhibit the development in these layers and thus modify the resulting images similarly to corrections by masking. As the density level corresponding to the "white" of the film is in general rather high, of a density of about 0.3 or 0.4, the interimage mechanisms act still below this level and thus yield highly saturated colors, so that the spectral absorptions of the real film dyes approach those of the so-called "block dyes" corresponding to optimal colors. The film and television domains then become very similar; besides this various electronic devices in the television channels also contribute to compensate the differences between film and television images (F. R. Clapper, private communication)

In the three diagrams of Fig. 3 the reproduction domains of the television screen and the projected color film, initially determined in log-luminance units and shown in dashed lines, were corrected for equivalent luminance and replotted in full lines. This adaptation of the reproduction limits to actual observer reaction demonstrates a very marked compensation in the tone scales of the inherent brightness differences of the various hues. Without taking in account equivalent luminances the brightness progressions of the saturated blue and yellow scales as well as of the saturated green and magenta scales would appear to differ considerably from the rather balanced tone scales of red and cyan hues. Corrected for equivalent luminance, however, the tone reproduction characteristics of the six principal hues of color reproduction are less different. The television screen scale from black to saturated blue, which uncorrected appears to increase in luminance ratio only by a factor of about two, covers after correction a luminance ratio of one to five, and the corresponding magenta scale increases from a luminance range of 1:7 to one extending from 1:15. The yellow and green scales, on the other hand, are only slightly modified by these corrections and those in the red-cyan plane remain almost unchanged. The corrected scales, showing the actual visual appearance of the tone progressions, thus make possible a quantitative analysis and comparison of the tone reproduction characteristics of television and film.

The evaluation of the reproduction quality of reflection prints can serve as another example of the usefulness of equivalent luminance information. In the inspection of images on white reflecting opaque support, obtained either photographically or by printing, the above discussed constancy of the equivalent luminance to nominal luminance ratio within the entire domain of photopic vision proves to be a very important factor. If these ratios were not constant over the whole range print appearance would change with illumination level; the parallelism of the equivalent luminance curves in the various luminance ranges, for instance in those determined by Hunt for print inspection, given above and shown by arrows in Fig. 3, however shows that widely varying conditions of print viewing must yield identical results because the saturation scales of colors appear to the observer identical at all levels.

THE GRAPHICAL REPRESENTATION OF COLOR REPRODUCTION CHARACTERISTICS

As mentioned before, all diagrams of Figs 1 and 3 are vertical planes of the photolorimetric grid shown in Fig. 2. The choice of a special cylindrical grid is closely related to the interpretation of the results of matching experiments in connection with photographic color reproduction in its widest sense. Graphical representations of such results require indeed a grid capable of including both surface colors and colored lights, and covering the whole domain from scotopic through mesopic to photopic vision. The CIE color spaces of 1931, 1960 and 1964, designed for the representation of the complete color gamut for only one state of general observer luminance adaptation, is neither adapted to this particular use, nor for the direct plotting of results in terms of optical densities which are the most frequently employed photometric quantities in investigations of photographic or related reproduction systems.

The new cylindrical space is based on the perceptually almost uniformly spaced CIE $u-v$ diagram and includes in addition the required luminance information. Its elements:

- 1960-CIE $u-v$ chromaticity diagrams of identical size in all horizontal plane cross-sections,
- a vertical luminance axis, laid through the locus u_0, v_0 of the chosen reference white and graduated logarithmically in standard units, i.e. in candelas per square meter,

yield in all vertical planes passing through this axis rectangular metric purity-luminance diagrams*. The scale ratio of the horizontal $u-v$ diagrams and the

* In the Barcelona meeting of the CIE in September 1971 "metric purity" was recommended as new colorimetric term for the purity measured in the $u-v$ diagram in analogy with $x-y$ excitation purity.

vertical log-luminance axis, determined by heterochromatic matches, is such that one unit of logarithm of luminance has the same length as one tenth of a unit of u or v in the horizontal diagrams (Kowaliski, 1967, 1969a). Besides its intrinsic advantage of extending from $-\infty$ to ∞ and of the possibility of the direct inclusion of density plots, the logarithmic luminance scale also has the advantage of being decimal and repetitive for every power of ten. This latter property is fundamental for demonstrating relationships such as those revealed by the equivalent luminance determinations over a wide range of luminance levels summarized above.

CONCLUSION

As the luminance scales of tone reproduction by photography or television include colors of all purity levels, ranging from white to the highest attainable saturations, they obviously depend on the equivalent luminances of colors. The evaluation of the colorimetric accuracy of color reproduction or of its limitations therefore requires a tridimensional grid containing the surfaces of constant photometric luminance, as well as a computational model for the interpolation of representative points lying between the planes and the surfaces of the grid. To this purpose the previously described measurements of equivalent luminances have been extended to rather high levels of luminance, reaching up to 3000 cd/m^2 for some colors of maximum metric purity. Applications of the grid and its usefulness are described for practical cases encountered in color reproduction.

REFERENCES

- Bartleson, C. J., and Breneman, E. J. (1967), Brightness reproduction in the photographic process. *Phot. Sci. Eng.* 11, 234
- Benson, K. B. (1970), Color television film viewing and transmission practices. *J. Soc. Motion Pict. Television Eng.* 79, 1085
- Brewer, W. L., Ladd, J. H., and Pinney, J. E. (1954), Brightness modification proposals for televising color film. *Proc. IRE* 42, 174
- Chapanis, A., and Halsey, R. M. (1955), Luminance of equally bright colors. *J. Opt. Soc. Amer.* 45, 1-6
- Draft USA Standard PH-2-32 (1968), Viewing conditions for the appraisal of color quality and color uniformity in the graphic arts industry
- Hunt, R. W. G. (1965), Luminance levels in colour transparencies and reflexion prints. *J. Photogr. Sci.* 13, 108
- Hunt, R. W. G. (1967), *The reproduction of colour*. 2nd ed. London, Fountain press, p. 381
- Judd, D. B. (1952), *Color in business, science and industry*. New York, Wiley, London, Chapman & Hall, p. 281
- Kowaliski, P. (1963), Tone reproduction in color scales. *J. Photogr. Sci.* 11, 169
- Kowaliski, P. (1963), Comité "Bases théoriques de la colorimétrie". *Actes VII^{me} Journées Int. de la Couleur*, Florence
- Kowaliski, P. (1967), Le problème de la luminance équivalente. *Rev. Opt.* 46, 359-369
- Kowaliski, P. (1969a), Equivalent luminances of colors. *J. Opt. Soc. Amer.* 59, 125-130

18 DISCUSSION

- Kowaliski, P. (1969b), Luminances équivalentes élevées des couleurs. Proc. 1st AIC congress "Color 69", Stockholm, 418-426
- Sanders, C. L., and Wyszecki, G. (1957), Correlate for lightness in terms of CIE-tristimulus values. I-II. J. Opt. Soc. Amer. 47, 398-404; 840-842
- Sanders, C. L., and Wyszecki, G. (1958), L/Y ratios in terms of CIE-chromaticity coordinates. J. Opt. Soc. Amer. 48, 228-232
- USA Standard PH 2.31-1969, Direct viewing of photographic color transparencies
- Wyszecki, G. (1967), Correlate for lightness in terms of CIE chromaticity coordinates and luminous reflectance. J. Opt. Soc. Amer. 57, 254-257

DISCUSSION

Bartleson: I wonder whether or not you considered different surround luminances in your measurements.

Kowaliski: In the initial measurements with comparatively low luminance levels (up to 200 cd/m²) several appropriate surround luminances, near those of the matching field luminance, were used. In the new instrument only one mean surround luminance is employed, but general room illuminance is adjusted according to the matching field luminance level.

Nimeroff: In the Brewer, Ladd, and Pinney paper they mentioned that they had the problem involved in three modes of illumination; the characteristics of a reflecting scene, transmittance of a photographically recorded and projected scene, and the emission of a scene by a television screen.

They dealt with only one part of the problem, the reflecting scene, One other problem is treated here by Kowaliski, the transparent scene. This leaves us with consideration of one more mode, the emission by a television screen.

Kowaliski: In the equivalent luminance determinations the appearance of the matching field can be that of a surface (reflection) colour when the surround luminance is higher than that of the field, but it is that of a transparency (transmitted light) with lower surround luminance. For this reason no further provision was made for this aspect of the problem.

19 SMALL AND MODERATE COLOR DIFFERENCES I VISUAL EVALUATION OF FMC-1 AND FMC-2 METRICS

F. W. BILLMEYER, Jr., E. D. CAMPBELL, G. L. KANDEL
AND J. MACMILLAN

*Rensselaer Polytechnic Institute, Troy,
New York 12181, U.S.A.*

INTRODUCTION

The need for extensive new visual observations of small and moderate color differences, and their relation to color differences calculated by new or established metrics, is unquestioned today. In initiating research in this area, we have undertaken two pilot studies specifically designed to determine which, if either, of the two metrics, FMC-1 (Friele 1961, MacAdam 1966, Chickering 1967) or FMC-2 (Wyszecki 1968, Chickering 1971), yields color differences corresponding more nearly to the way such differences are seen by observers. That is, we wished to answer such questions as: Are the factors k_1 (weighting ΔC) and k_2 (weighting ΔL) by which FMC-2 differs from FMC-1 necessary? Does their inclusion generate color-difference values which are more nearly in accord with visual assessment?

In passing, we note that although the FMC metrics are designed to have the properties of a *ratio scale* and are based on a unit of difference (the MacAdam, 1942, unit) which is at the *threshold*, i.e. not readily perceptible, our experiments are attempts to determine which equation provides a more adequate *interval scale* for *suprathreshold*, i.e. readily perceptible, differences.

In undertaking the larger research, including the pilot studies here reported in this field, we have attempted to adhere as closely as possible to conditions found in typical industrial applications of color-difference judgment. Thus, our samples are prepared colored surfaces rather than lights (surface rather than aperture mode of viewing). We have chosen these because we feel that judgments of surface colors are more pertinent to practical situations. While this choice has one advantage, in that it limits the extent of the (visual) color

space with which we must ultimately deal, it also has the disadvantage that continuously variable stimuli are not conveniently at our disposal for our psychophysical measurements. Our stimuli must be carefully selected, and too often our experiments must be tailored to the particular stimulus values which happen to be available among our samples.

EXPERIMENTAL

1. *Samples.* We are indebted to Mrs. Dorothy I. Morley (Metal Box Co., Ltd., Acton, England) for making available several sets of samples prepared by lithographic printing on tinplate, followed by varnishing. Each set consisted of approximately 30 samples clustered in a small region of color space. In this work, we utilized sets in the regions with ISCC-NBS color names vivid blue (approximately $x=0.18$, $y=0.16$, $Y=6$), vivid yellow ($x=0.45$, $y=0.49$, $Y=50$), deep purplish red ($x=0.48$, $y=0.25$, $Y=7$), and vivid yellowish green ($x=0.31$, $y=0.51$, $Y=29$).

2. *Visual Observing Conditions.* Samples were presented for observation in pairs, with a minimum dividing line between the two, against a neutral gray surround of about half the luminance of the samples. The samples subtended a field of about 2° at normal viewing distance. Illumination was semi-diffuse using Macbeth "Norlite" fluorescent tubes simulating CIE Illuminant D_{65} . The observer was allowed to adjust the sample position, if necessary, to avoid specular reflections; viewing was usually near the normal. Care was taken that the entire surrounding field was neutral gray.

3. *Measurements.* The samples were measured on both a Kollmorgen KCS-40 abridged spectrophotometer and a Zeiss DMC-25 spectrophotometer; corresponding results from the two instruments were in good agreement. The color differences reported were calculated from KCS-40 measurements, using diffuse illumination by an illuminant A source and near-normal viewing with specular component excluded. Reflectances relative to pressed BaSO_4 at 10 nm intervals from 380 to 750 nm were integrated by the weighted-ordinate method to tristimulus values for CIE Illuminant D_{65} and the 1931 CIE 2° Standard Observer. FMC-1 and FMC-2 color differences were calculated by digital computer. Care was taken to measure the same portion of the sample used in the visual observations. Three measurements were made on different portions of the viewed area, and the average of the three sets of tristimulus values was used for the color difference calculations. Table I gives the average tristimulus values of each sample.

Table I Tristimulus Values of Samples, for CIE Illuminant D_{65} and the 1931 Standard Observer.

Sample	X	Y	Z	Sample	X	Y	Z
M73	12.64	6.83	7.03	Y39	50.26	50.33	4.18
M82	12.62	6.85	6.85	Y25	46.90	48.21	4.52
M71	14.19	7.11	8.08	Y42	46.71	48.84	5.16
M87	14.29	7.28	7.78	Y28	46.92	49.51	6.57
M89	14.72	7.49	7.51	Y29	47.07	50.07	6.77
M93	16.03	8.09	8.27	Y2	47.62	49.94	4.05
M88	14.83	7.54	8.21	Y16	47.16	50.09	5.99
M94	17.12	8.69	9.60	Y44	47.07	50.03	5.60
M14	16.80	8.56	7.38	Y48	47.61	50.63	5.79
M56	17.24	8.20	10.77	Y19	46.47	48.60	4.95
M83	13.24	7.29	6.77	Y47	46.93	49.56	4.81
M86	14.04	7.21	8.68	Y32	45.89	48.23	6.26
M80	13.05	7.05	9.11	Y38	46.04	47.90	4.98
M81	12.90	7.09	7.73	Y1	44.85	49.27	5.39
M70	14.60	7.00	7.31	Y6	46.17	46.17	5.49
M90	14.50	7.37	7.08	G1	16.34	28.20	9.63
M67	16.23	7.88	9.37	G3	16.84	28.04	10.67
M100	16.19	8.16	9.28	G5	18.47	29.42	11.71
B19	8.18	8.14	25.99	G7	17.06	28.22	9.77
B23	7.21	7.20	22.57	G8	16.69	27.91	10.00
B7	6.86	6.36	24.26	G9	18.77	29.92	11.70
B13	6.49	6.04	22.70	G10	16.58	28.28	10.01
B11	5.95	5.73	20.04	G11	16.91	28.60	9.60
B31	6.58	6.43	21.96	G12	17.26	28.93	9.06
B15	5.86	5.97	19.76	G13	18.20	28.30	10.18
B26	6.66	6.29	19.76	G14	18.98	28.14	11.20
B10	5.60	5.68	18.90	G15	18.30	29.91	10.55
B29	6.54	5.71	22.59	G16	18.32	28.57	11.52
B24	5.88	5.49	20.52	G17	18.00	28.37	10.21
B25	6.22	5.38	21.37	G18	18.27	28.94	10.36
B1	6.69	5.84	24.75	G19	19.21	29.96	11.64
B30	6.51	6.02	22.62	G20	19.18	28.63	12.96
B9	5.88	5.71	20.46	G24	18.46	30.10	10.97
B14	6.26	5.70	22.37	G25	17.25	27.91	9.44
B4	6.75	6.06	24.62	G26	17.73	28.59	9.40
B12	6.61	5.95	24.03	G28	18.02	28.47	10.65
Y34	47.64	49.71	4.84	G29	17.13	29.57	10.07
Y45	46.97	48.90	4.84	G30	17.34	29.59	9.90
Y40	48.24	50.06	6.20	G31	17.51	28.88	10.47

FIRST EXPERIMENT

This experiment was devised and carried out by E. D. Campbell. Samples in the vivid blue, vivid yellow, and deep purplish-red regions were used: the vivid yellowish-green samples were used only in the second experiment.

1. *Sample Selection.* For the colors used in this experiment, some pairs of samples (designated C-pairs) were selected to have ΔL negligible compared to ΔC , while others (designated L-pairs) were selected to have ΔC negligible compared to ΔL . Only in one or two instances, where ΔE was of the order of one

unit, were there slight exceptions to this rule. This is not considered serious since we expected that these pairs would invariably be judged "smaller than" those to which they were compared.

2. *Observations.* The tests were conducted using the well-known frequency method. Two observers were used, both male, ages 24 and 25, both shown to have normal color vision by the AO-HRR, Ishihara and Farnsworth-Munsell 100 Hue tests, and both experienced in judging small color differences. One C-pair and one L-pair were shown at a time to the observer, who indicated whether the C-pair appeared to have a color difference greater or less than that of the L-pair. Observations were continued, C-pairs being presented in random order, until 5 judgments of each C-pair vs. each L-pair had been made. Experiments with the three different colors were intermixed in random order. In all, the two observers made some 4600 judgments,

3. *Treatment of Data.* The objective of this experiment was to calculate the point of subjective equality between differences in lightness and in chromaticity, and to compare this with the predictions of the FMC-1 and FMC-2 metrics. This point was determined from the observations in two ways:

First, the fraction of the times a given C-pair was judged to have a greater (or lesser) color difference than a given L-pair was plotted against ΔE as calculated using one or the other metric, and the value of ΔE for which that fraction equaled 0.50 was estimated graphically. Fig. 1 shows such a plot, in which the data fit somewhat better than average. Note that the goodness of fit of the

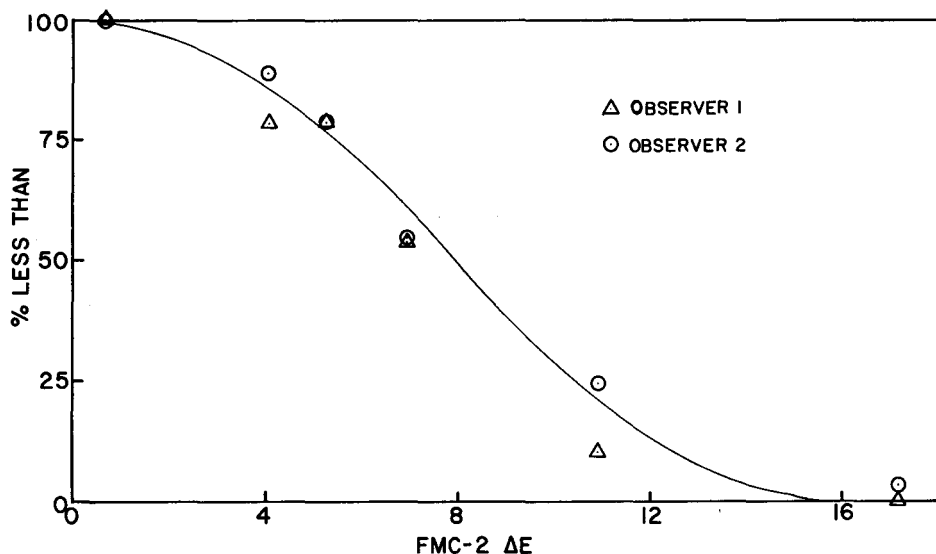


Fig. 1. Visual judgments of vivid blue BC pairs plotted against FMC-2 color differences, for pair BL3; data representing a good fit.

data does not necessarily imply that the FMC-2 metric is better than FMC-1, but does show that the observations are consistent with the shape of FMC perceptibility ellipse at that point in color space. In some cases, however, the visual data do not lie on a smooth curve. We believe this implies that these data are not consistent with the shape of the FMC perceptibility ellipse at this point. This is typical of results with the deep purplish-red samples, as discussed below.

The second method of determining the point of subjective equality is based on our observation that, at least in the cases where the data fit on a smooth curve, they yield a linear plot when the fraction of the times a C-pair was judged greater (or less) is plotted on normal probability paper *versus* a linear scale of ΔE . For such plots, the best line of fit, the point of 50% probability (subjective equality), and the standard deviation of this point, could be computed by a least-squares calculation. Fig. 2 shows one of these plots.

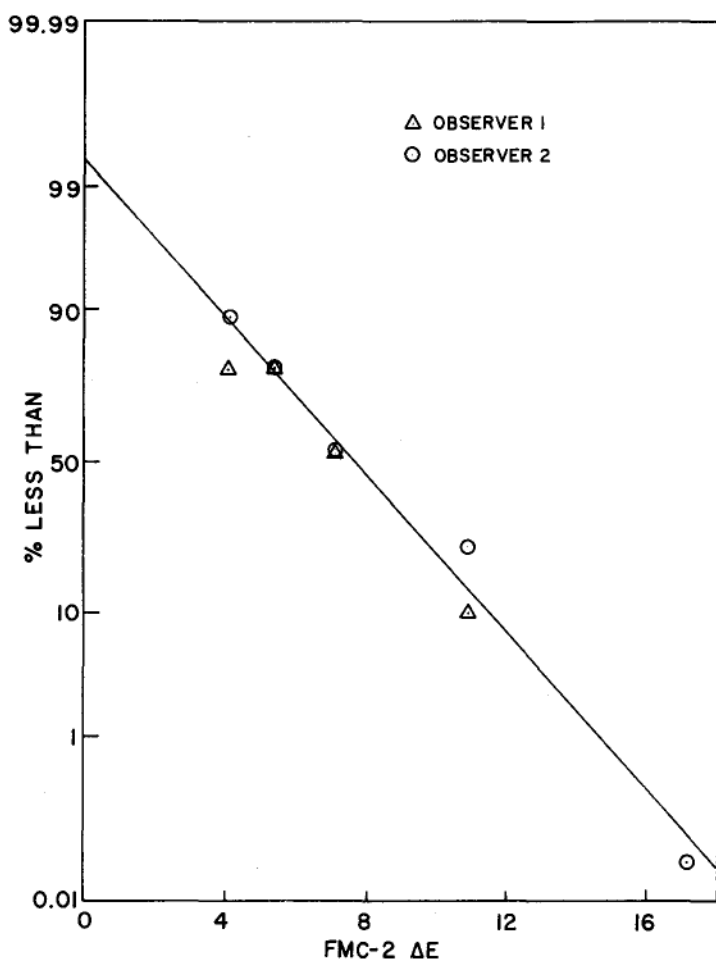


Fig. 2. Plot of the data of Fig. 1 on normal probability paper

4. *Results.* The results of the visual observations are summarized in Tables II-IV which list, for each combination of L-pair and C-pair, the fraction of the times the C-pair was judged to have a smaller color difference than the

Table II Per Cent of Time C-Pair Judged to Have Color Difference Less than that of L-Pair, for Vivid Blue Sets. Data for Two Observers Averaged.

	L-Pair Nos.		B7-B13	B11-B31	B19-B23
	FMC-1 ΔE		4.90	11.00	10.65
	FMC-2 ΔE		1.85	3.65	4.15
C-Pair Nos.	FMC-1 ΔE	FMC-2 ΔE	Av. % Less	Av. % Less	Av. % Less
B4 -B12	1.76	0.63	86	99	100
B9 -B14	5.09	4.09	66	80	84
B1 -B30	6.68	5.38	39	69	79
B24-B25	8.69	6.93	4	35	54
B10-B29	13.65	10.95	0	11	17
B15-B26	21.15	17.17	0	2	2

Table III Per Cent of Time C-Pair Judged to Have Color Difference Less than that of L-Pair, for Vivid Yellow Sets. Data for Two Observers Averaged.

	L-Pair Nos.		Y34-Y45	Y1-Y6
	FMC-1 ΔE		1.68	2.88
	FMC-2 ΔE		1.43	2.28
C-Pair Nos.	FMC-1 ΔE	FMC-2 ΔE	Av. % Less	Av. % Less
Y44-Y48	1.26	1.06	95	55
Y28-Y29	2.08	3.26	95	37
Y19-Y47	2.64	3.58	77	20
Y32-Y38	3.90	6.74	50	11
Y25-Y42	4.34	7.40	0	0
Y2 -Y16	5.39	9.51	55	0
Y39-Y40	9.76	17.18	0	0

Table IV Per Cent of Time C-Pair Judged to Have Color Difference Less than that of L-Pair, for Deep Purplish-Red Sets. Data for Two Observers Averaged.

	L-Pair Nos.		M89-M93	M88-M94
	FMC-1 ΔE		6.68	11.72
	FMC-2 ΔE		2.71	4.19
C-Pair Nos.	FMC-1 ΔE	FMC-2 ΔE	Av. % Less	Av. % Less
M73-M82	1.31	1.12	63	70
M71-M87	3.43	2.82	99	99
M67-M100	6.67	5.93	56	84
M80-M81	7.90	6.85	72	76
M70-M90	10.16	8.73	74	93
M83-M86	14.78	12.89	14	26
M14-M56	21.81	20.13	2	8

Table V Color Difference at the Point of Subjective Equality of C-Pairs and L-Pairs and its Standard Deviation, as Calculated from the Least Squares Treatment, Compared to Instrumentally-Measured Color Differences for L-Pairs.

	FMC-1			FMC-2		
	Calcd.	Experimental		Calcd.	Experimental	
Standard L-Pair		Obs. 1	Obs. 2		Obs. 1	Obs. 2
B7 -B13	4.90	3.7 ± 1.6	6.4 ± 1.9	1.85	2.1 ± 1.8	4.9 ± 1.7
B11-B31	11.00	7.3 ± 3.4	12.6 ± 3.4	3.61	5.6 ± 2.9	10.2 ± 2.8
B19-B23	10.94	9.6 ± 2.0	11.9 ± 3.8	4.15	7.6 ± 1.7	9.5 ± 3.1
Y34-Y45	1.68	2.9 ± 1.5	4.3 ± 1.2	1.43	4.4 ± 2.8	7.2 ± 2.2
Y1 -Y6	2.88	1.8 ± 0.9	0.0 ± 2.2	2.28	2.4 ± 1.9	-0.9 ± 4.2
M89-M93	6.68	7.1 ± 3.8	14.2 ± 5.1	2.71	6.2 ± 3.5	12.5 ± 4.4
M88-M94	11.72	9.8 ± 3.4	17.7 ± 8.6	4.19	8.7 ± 3.1	15.8 ± 7.7

L-pair. All further discussion of these data is based on the least-squares treatment listed above as the second method of analysis.

The results of the least-squares analyses are summarized in Table V. Inspection of this Table shows that, for the case of vivid blue, the color difference calculated by the FMC-1 metric falls within one standard deviation of the experimentally observed point of subjective equality of ΔC and ΔL as determined by the least squares method. In the case of the FMC-2 predictions, this is true in only two of the six cases (three L-pairs, two observers), and in those two the standard deviations are quite large compared to the ΔE at that point. We conclude that FMC-1 represents our visual data quantitatively, while FMC-2 fails to do so by a wide margin (At this lightness level, k_1 is approximately 0.82 and k_2 approximately 0.13).

In the case of vivid yellow, the visual data fall on relatively smooth curves, but inspection of Table V shows that good quantitative agreement is not obtained with either color-difference metric (Here k_1 is about 1.78, and k_2 about 0.77).

Still worse, the visual data for the deep purplish-red samples do not at all fall on smooth curves. Inspection of the data for individual C-pairs shows clear discrepancies from the predictions of either FMC-1 or FMC-2. Neither metric is at all representative of the visual results at this point in color space (here k_1 and k_2 are about the same as for vivid blue).

As a further elucidation of Table V, we have plotted in Figs 3 and 4 the value of ΔE for the C-pairs at the point of visual equality and its standard deviation (indicated by the length of the vertical line), against the calculated color difference for the L-pair used in the comparison. In these graphs a 45°

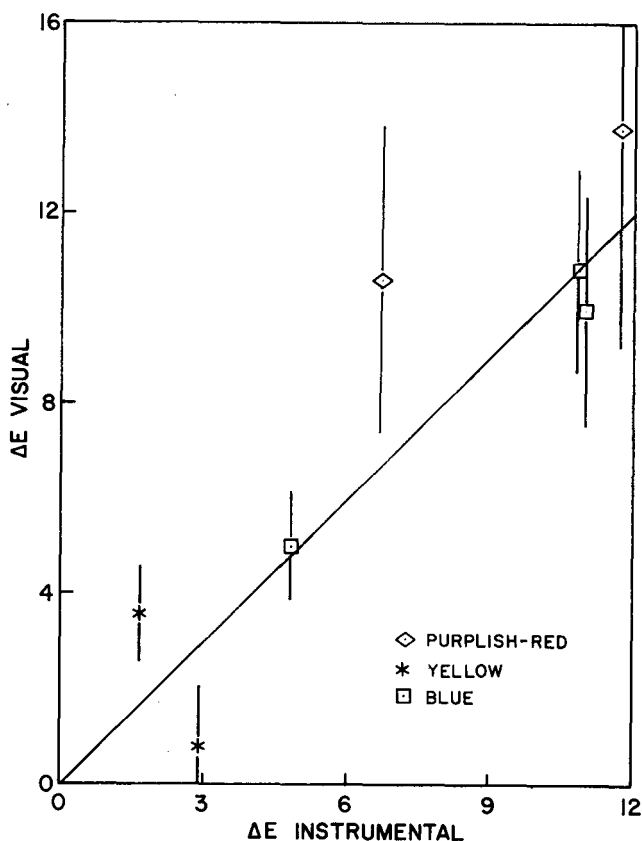


Fig. 3. Points of visual equality of C-pairs to a given L-pair, and their standard deviations, plotted against the color difference of the L-pair, calculated with the FMC-1 metric. Data for two observers averaged.

line is included to indicate the point of exact agreement. The superiority of the fit for FMC-1 (Fig. 3) over that for FMC-2 (Fig. 4) is clear.

In the case of vivid blue, for which the preference for FMC-1 over FMC-2 was clear, it was possible to analyze the data further to determine the extent to which the value of k_1 and of k_2 contributed to the lack of agreement with FMC-2.

For FMC-1, $\Delta E_1 = [\Delta C^2 + \Delta L^2]^{\frac{1}{2}}$, while for FMC-2, $\Delta E_2 = [(k_1 \Delta C)^2 + (k_2 \Delta L)^2]^{\frac{1}{2}}$. Since ΔL is negligibly small for our C-pairs, and ΔC negligibly small for our L-pairs, we may write $\Delta E_{1,C} = \Delta C$, $\Delta E_{2,C} = k_1 \Delta C$, $\Delta E_{1,L} = \Delta L$, and $\Delta E_{2,L} = k_2 \Delta L$. If, at the point of visual equality, we compare values of $\Delta E_{1,L}$ and $\Delta E_{2,C}$, we obtain a measure of the effect of k_1 , whereas if we compare values of $\Delta E_{1,L}$ and $\Delta E_{2,L}$, the effect of k_2 is seen.

Our results show that there is agreement 67-70% of the time when k_1 is

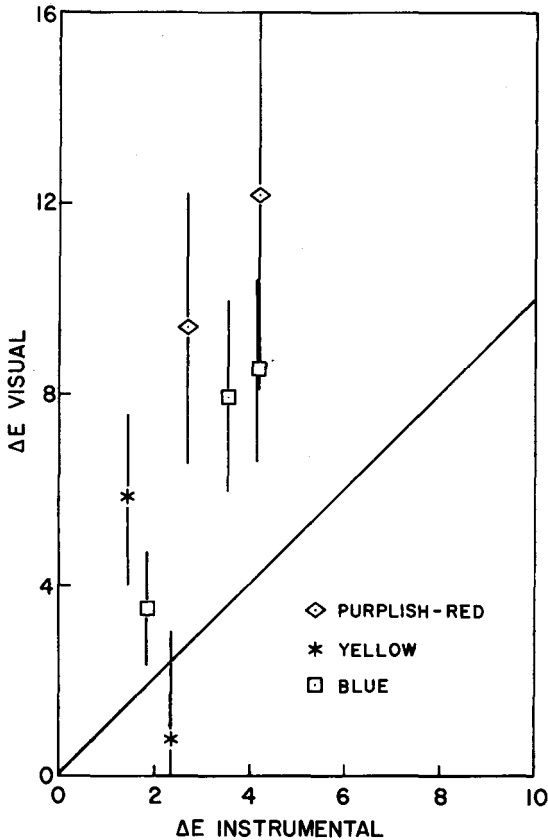


Fig. 4. Same as Figure 3, calculated with the FMC-2 metric.

being tested, but only 16-30% of the time when k_2 is being tested. These conclusions are amplified by the graphs of Figs 5 and 6, which are of the type of Figs 3 and 4, previously discussed. It seems clear that, in this region of color space, FMC-2 is overcorrecting lightness differences, so that values of ΔE calculated for the L-pairs are numerically too small.

SECOND EXPERIMENT

In this experiment, devised and carried out by J. MacMillan, only samples in the vivid yellowish-green region were used.

1. *Sample Selection.* The selection of sample pairs for the Second Experiment was made without regard to the components of the color differences. Instead, pairs were selected for which the values of ΔE calculated by the FMC-1 and FMC-2 metrics were rather different; these are designated D-pairs. For

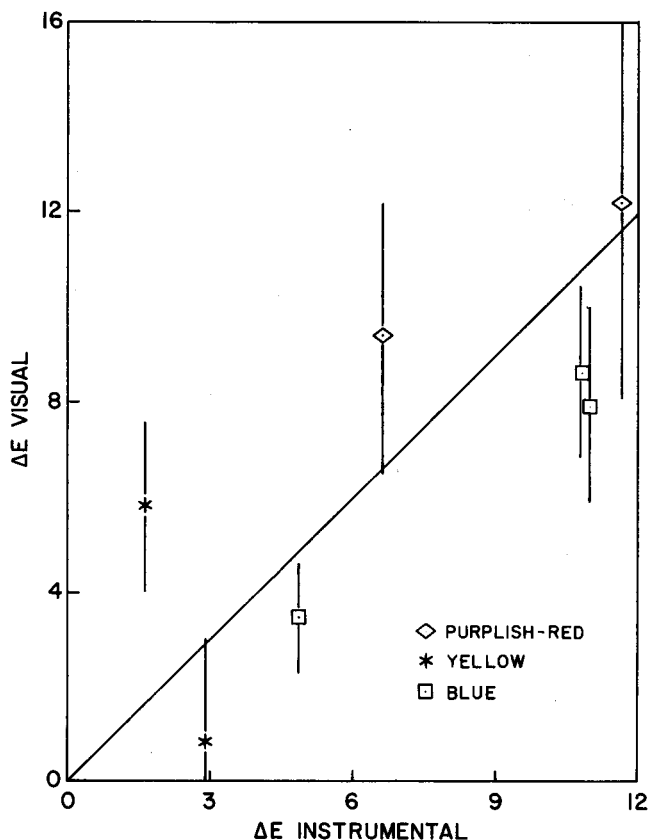


Fig. 5. Same as Fig. 3, plotted to show the effect of k_1 (ΔE for C-pairs using FMC-1; and for L-pairs, FMC-2).

each of the D-pairs, a second pair of samples was selected, for which the two values of ΔE were rather similar, and fell between the two values calculated for the D-pair; the second pair of each set we designate the S-pair. A visual judgment that the color difference of the D-pair is greater than that of the S-pair favors FMC-2; the reverse judgment favors FMC-1; and the judgment that the two pairs of the set have the same color difference favors neither metric. The color differences for the D-pair and S-pair of each set are given in Table VI.

2. *Observations.* Fifty observers were used in this experiment, about 1/5 being female and 4/5 male. Ages ranged from 18 to 65, and experience in judging color differences from none to extensive. All were screened for normal color vision (AO-HRR or Ishihara test) and to be sure they understood the instructions.

Each observer judged the nine sample sets, presented in random order, at least three times. The color difference of one pair from a set could be judged

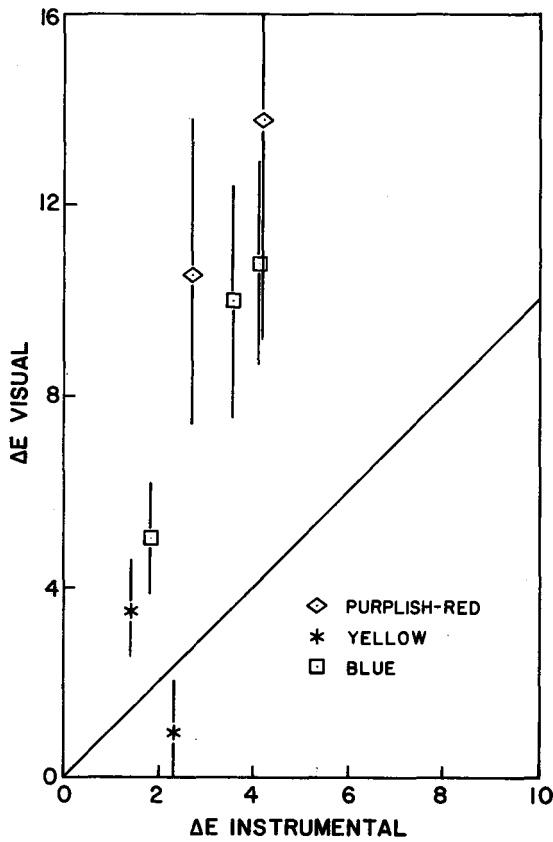


Fig. 6. Same as Fig. 3, plotted to show the effect of k_2 (ΔE for C-pairs calculated using FMC-2; and L-pairs, FMC-1).

Table VI Green Colored Pairs and their Color Differences.

Expt. Designation	Spl. #	Spl. #	FMC-1			FMC-2		
			ΔC	ΔL	ΔE	ΔC	ΔL	ΔE
S 1	G18	G19	2.58	3.57	4.30	3.79	2.25	4.28
S 2	G9	G17	3.42	5.12	6.41	5.08	3.26	6.20
S 3	G15	G26	3.15	4.34	5.56	4.68	2.76	5.64
S 4	G5	G16	2.70	2.77	3.97	3.98	1.75	4.44
S 5	G24	G24	3.79	5.18	6.69	5.64	3.31	6.82
S 6	G3	G30	4.06	5.12	6.34	5.89	3.19	6.52
S 7	G8	G11	1.85	2.29	2.90	2.67	1.42	3.00
D 1	G7	G10	4.37	0.24	4.40	6.35	0.15	6.39
D 2	G13	G25	4.91	1.49	5.18	7.14	0.93	7.27
D 3	G12	G31	3.76	0.02	3.76	5.52	0.01	5.52
D 4	G14	G20	5.22	1.88	5.40	7.47	1.17	7.58
D 5	G1	G29	0.03	4.54	4.64	0.04	2.86	2.90

larger than, smaller than, or equal to that of the other pair. The first observation on a given set, though recorded, was not counted; since many of the observers were completely naive, it was felt that a practice period was desirable. After this group of observations, sets were presented until the observer had made the same judgment on that set two times out of three (or rarely, three times out of five). On the average, an observer required only five extra observations over the minimum 27 to achieve this frequency. Thus, some 1600 observations in all were made.

3. *Treatment of Data.* The data from the Second Experiment were evaluated as follows:

First, totals were obtained of the number of judgments favoring FMC-1, those favoring FMC-2, and those selecting the third choice corresponding to neither of these. These totals are listed, together with identification of the pairs compared, in Table VII.

Table VII Number of Observers Selecting Each Possible Choice for Second Experiment.

Pair Desig.		FMC-1	FMC-2	Neither
S1	D1	7	38	3
S2	D1	38	4	6
S3	D1	36	6	5
S3	D2	2	13	33
S3	D3	29	5	15
S4	D3	0	10	39
S5	D4	29	17	1
S6	D4	36	11	3
S7	D5	15	26	5
Totals		192	130	110

Second, the data were subjected to several statistical tests, which will not be described in detail, although the places at which they reinforced the conclusions to be drawn will be pointed out.

4. *Results.* It may first be mentioned that several chi-square "goodness-of-fit" tests were made which demonstrated that the observers were definitely not guessing. In terms of set judgments, out of 450 possible selections (50 observers, 9 sets each), 192 judgments favored FMC-1 calculations as correct, 130 favored FMC-2, 110 corresponded to the third choice, and 18 were rejected for various reasons. In terms of individual sets, a majority of the observers (the *smallest* majority being 26 out of 46) selected the judgment corresponding to the correctness of FMC-1 for five sets, that for FMC-2 for two sets, and the third choice for the remaining two.

The statistical tests showed, however, that the slight bias in favor of FMC-1 was not significant. Application of the t-test showed that none of the following

hypotheses could be rejected: FMC-1 and FMC-2 predict the visual results equally well; the FMC metric predicts the results equally well with $k_1 = k_2 = 1$ (FMC-1) and with other values for these factors; and the values of these factors used in FMC-2 ($k_1 = 1.57$ and $k_2 = 0.62$) predict the results as well as would use of another set of factors making the "third choice" always correct.

A further analysis of the data suggests that a set of factors could in principle be found experimentally which would, for this point in color space, always predict correctly the *direction* of the difference in ΔE 's. In this set, k_1 should be similar to that now used in FMC-2, but k_2 should be nearer unity.

GENERAL CONCLUSIONS

1. Despite some not unexpected difficulties, the use of surface-colored samples does not preclude obtaining significant results in the visual judgments.
2. In some regions of color space, neither FMC-1 nor FMC-2 represent the visual data well enough to allow any meaningful conclusions to be drawn as to which is better.
3. In the one region (vivid blue) where one of these metrics was quantitatively substantiated by the visual observations, it was FMC-1 which fitted the visual data.

ACKNOWLEDGMENTS

The research to which these pilot studies were preliminary is supported in part by the Munsell Color Foundation, The Sherwin-Williams Company, the Dry Color Manufacturers' Association, and the National Science Foundation. This is Contribution No. 42 from The Rensselaer Color Measurement Laboratory.

REFERENCES

- Chickering, K. D. (1967) Optimalization of the MacAdam Modified 1965 Friele Color Difference Formula. *J. Opt. Soc. Amer.* 57, 537-541
- Chickering, K. D. (1971) FMC Color Difference Formulas. Clarification Concerning Usage. *J. Opt. Soc. Amer.* 61, 118-122
- Friele, L. F. C. (1965) Further Analysis of Color Discrimination Data. *J. Opt. Soc. Amer.* 55, 1314-1319
- MacAdam, D. L. (1942) Visual Sensitivities to Color Differences in Daylight. *J. Opt. Soc. Amer.* 32, 247-274
- MacAdam, D. L. (1966) Smoothed Versions of Friele's 1965 Approximations for Color Metric Coefficients. *J. Opt. Soc. Amer.* 56, 1784-1785
- Wyszecki, G. (1968) Recent Agreements Reached by the Colorimetry Committee of the Commission Internationale de l'Eclairage. *J. Opt. Soc. Amer.* 58, 290-292

DISCUSSION

Friele: In Table II the calculated order of Y34-Y45 and Y1-Y6 is reversed compared to the observed differences. As both show mainly a lightness difference, one would expect that any colour difference formula would order them in the correct order.

Billmeyer: You are correct that there is a reversal, but the data clearly substantiate this and the results with two observers are consistent. The color differences are near threshold, however, and we can only conclude that something, perhaps surface defects on the samples, has biased either the visual or the instrumental results. The second part of your question is true but is not relevant since we never compared pair Y34-Y45 to pair Y1-Y6, but always compared an L-pair to a C-pair.

Ganz: 1. Do you intend to test other colour difference formulae than FMC-1 and FMC-2?

2. Did you also include neutral pairs as a reference colour difference, as are used for the visual assessment of fastness tests?

Billmeyer: 1. We expect to do this in extensions of this work to be published elsewhere at a later date.

2. We did not include neutral pairs, using the L-pairs as reference color differences in the First Experiment. The Second Experiment did not require a fixed reference color difference. Furthermore, we never asked for ratio judgments, but only asked which of two pairs, seen simultaneously, had the larger color difference.

Brookes: How can you differentiate between the influence of k_1 and k_2 (p. 218)? Your experiment compares only the size of lightness differences with that of chromaticity differences, so that a common factor cancels out.

The experimental point of visual equality is defined in the paper by the statement that ΔE (chromaticity) = ΔE (lightness). You use then the experimental tristimulus values of such a pair of samples to test, whether the formula FMC-1 ($\Delta C = \Delta L$) or FMC-2 ($k_1 \Delta C = k_2 \Delta L$) agrees better with this condition. The latter equation, written in the form $(k_1/k_2) \Delta C = \Delta L$, shows clearly, that only the ratio k_1/k_2 is tested and compared with the FMC-1 condition where $k_1/k_2 = 1$.

In your own words :You use the comparisons $\Delta L = k_1 \Delta C$ and $\Delta C = k_2 \Delta L$. To fulfill the experimental condition as stated above, one can either e.g. in the first of these equations change k_1 or add a factor k_2 on the left side, because no information is present which side of the equation disagrees with the experiment. Again, only the ratio is determined for k_1 and k_2 .

Billmeyer: The first part of your remarks is correct. When we compare to the visual results both ΔC and ΔL calculated with FMC-2 metric, we do obtain only information about the ratio k_1/k_2 . However, we then compare the visual

results ΔC calculated by FMC-1 and ΔL calculated by FMC-2, for example. This is equivalent to calculating ΔE for both the C-pair and the L-pair by a new metric in which $k_1 = 1$ and k_2 has the value required by FMC-2. We do not see how any ratio can be involved here. Similarly, when we compare the visual results ΔC calculated by FMC-2 and ΔL by FMC-1, the situation is reversed and we test a metric having k_1 from FMC-2 and $k_2 = 1$.

Clarke: Are your samples large enough and/or uniform enough to permit valid 10° subtense observation?

Billmeyer: They are large enough, and the uniformity is not too bad. Dr. Day of our laboratory expects to compare observations at 2° and 10° subtense for these samples, but they may not show enough metamerism to make this a valuable experiment.

Hunt: I would like to ask a question about the illuminant, which was a Macbeth "Norlite" fluorescent "D65" tube. The computations however were based on CIE D_{65} . Do you intend to rework the data.

Billmeyer: Probably not, since these were only pilot experiments. We do recognize the need to use the exact spectra power distribution in the integrations, and will certainly do so as the work is expanded.

Wyszecki: I would like to emphasize that the experiment conducted is a "perceptibility" study which is highly commendable. There are only few such experiments as compared to "acceptability" studies. In general, acceptability data do not test a color-difference formula because color-difference formulae are exclusively formulae designed to predict perceptibility data. In many industrial applications a locus of constant acceptability around a given standard color is *not* identical to a locus of constant perceptibility.

Jaekel: If Dr. Wyszecki meant that perceptibility is *more* important than acceptability, I should like to disagree with him strongly, without going so far as to say that perceptibility is unimportant. However, in industrial situations, acceptability matters far more.

Wyszecki: I did not intend anything like that at all. I only thought that the FMC formulae are developed to predict perceptibility and from that point of view these experimental results provide indeed a fair test of the formulae.

McLaren: I believe that where the important judgement is made by consumers, there is no difference between perceptibility and acceptability: if the difference is readily perceived the goods are likely to be rejected.

20 COLOUR TOLERANCES IN THE PAINT INDUSTRY

I. G. H. ISHAK AND S. ROYLANCE

*Paint Research Association
Teddington, England*

INTRODUCTION

Reproducible production of colours within specified tolerances is of extreme importance in the paint industry, in particular, for industrial finishes. At present this relies mainly upon the visual judgement of a "colour matcher" who is a highly experienced observer. However, disagreements between matchers are not infrequent, and sometimes the same observer can give conflicting judgements on different occasions. An objective and reliable method of colour matching is needed to reduce the number of customer complaints.

For colour measurement and specification there are two requirements which need to be fulfilled: due to variations in the colour vision characteristics of observers, a standard observer is needed; and a system for specifying colour differences in agreement with visual perception is required. Unfortunately for the present only a compromise is possible on both of these issues.

The 1931 CIE 2° standard observer is not a true representative average. At the 1955 CIE conference, modifications to the standard observer were proposed by Judd, but, for practical reasons, it was decided to make no change to the 1931 observer. Apart from differences in the judgement of individuals, the response of every colour measuring instrument differs from the CIE observer. Also it has been reported in several investigations (Robertson and Wright, 1965; and Billmeyer, 1965) that measurements on the same sample, using different instruments of the same make, do not agree satisfactorily.

Many attempts have been made to produce a system with uniform scaling (for example, Judd, 1939; Hunter, 1942; Bond and Nickerson, 1942; MacAdam, 1943; Richter, 1955; Friele, 1959; and Wyszecki, 1963). Comparisons between the systems have been made by several workers, (McLaren, 1969), and several papers are to be presented at this symposium on this topic, (Simon, Jaeckel, and Dinsdale and Malkin). The relative merits of the various systems

can be judged using two criteria: the projection of the CIE System should be uniform on the basis of the work by Wright (1941) and MacAdam (1942) on visual discrimination; secondly, the Munsell system should be uniformly represented on these scales. The systems are all more uniform than the CIE in these two respects. Thurner and Walther (1969), however, conclude from comparisons made between various colour difference formulae that there is no formula which gives an entirely satisfactory correlation with visual assessment. Several systems are in use at present in each industry, resulting in confusion. Since there is no evidence of the superiority of one colour difference formula over the others, adoption of the CIE recommendations could solve the problem partly, until sufficient evidence is produced in favour of an alternative system. Until one colour system is adopted throughout the paint industry, it will be impossible to establish general tolerances.

One of the problems in determining a uniform colour space is that we do not have sufficient knowledge on psychological scaling. There is no evidence that equal perceptual differences will sum to give a uniform scale throughout the colour gamut. The Munsell system was based on equal perceptual differences, however, Ishak *et al* found a failure of Munsell chroma to be proportional to estimates of saturation. This was confirmed by work carried out by Judd and Nickerson (private communication by Judd). It is doubtful whether the problem of a uniform colour system will be solved in the near future, at least, not until more is known about psychological scales.

The purpose of this work is to recommend a practical method for tolerance specification for industrial finishes. Samples differing in colour from a standard by an amount judged by experienced matchers to be within the acceptable tolerance for industrial finish were measured on the Fibre Optics Colorimeter (Ishak, 1970), the Colormaster, and the Spectronic 505. The colour difference between each sample and the standard was computed in four colour difference units, and the results from the three instruments were compared.

METHODS

From the files of five firms, standards from the production ranges were chosen, together with several samples which had been passed as within the acceptable tolerance for industrial finish by the manufacturer, and had been accepted by the customer. This method for investigating tolerances is more practical than the situation where an observer is given two panels and must judge whether or not their colour difference is greater than the acceptable tolerance.

Thirty six standard colours were used in this study. Each standard and its corresponding samples were measured on the three instruments. ΔX , ΔY , and ΔZ were found from the readings from each instrument, and ΔE between each

sample and the standard was computed in the following colour difference units: NBS, CIE, Friele-MacAdam-Chickering (FMC-2), and Cube-Root.

RESULTS AND DISCUSSION

Table I shows the X , Y , Z specification for each of the 36 standards as computed from spectrophotometric data from the Spectronic 505, and the energy distribution of illuminant C.

Table II shows the range of ΔE values obtained from each instrument for the samples of each paint. Maximum and minimum values of ΔE are given, each in four colour difference units.

For purposes of comparison, the samples in each set were graded according to their colour difference from the standard. This order did not vary significantly from one colour system to another. It was observed that, in general, the magnitudes of ΔE in NBS, CIE and Cube Root units are approximately the same, while ΔE in FMC-2 units is 2-3 times as large. On comparing the order of gradation of each set of panels for the three instruments, no correlation was found, except for the lightest colours.

For the lighter colours—whites and beiges (e.g. paints 11, 12, 15, 21, 30 and 36), ΔE from all three instruments was generally lower. Some of the dark green panels (paints 6, 7, 17) showed an exceptionally high ΔE when measured on the Fibre Optics Colorimeter. For some dark reds (paints 9, 13, 18, 26, 35), the Colormaster gave higher values of ΔE than the other two instruments.

Table I Tristimulus Values of Standards.

Paint	X	Y	Z	Paint	X	Y	Z
19.	57.89	64.89	39.65	1.	14.76	15.54	17.55
20.	35.28	43.76	59.28	2.	24.63	26.70	33.88
21.	37.54	37.56	32.11	3.	1.58	1.74	4.27
22.	2.06	2.53	7.78	4.	6.10	6.00	6.52
23.	19.84	22.51	16.48	5.	70.58	72.79	74.73
24.	0.80	0.72	1.81	6.	21.49	24.02	24.25
25.	14.70	14.87	12.67	7.	1.53	1.78	2.09
26.	10.32	5.27	1.35	8.	2.36	1.45	1.07
27.	60.66	64.59	47.88	9.	12.19	6.41	1.94
28.	9.96	14.05	18.97	10.	1.08	1.41	1.35
29.	2.00	2.18	4.49	11.	65.13	67.07	69.94
30.	65.92	69.62	68.74	12.	17.84	18.83	22.61
31.	2.02	2.92	2.61	13.	15.42	8.25	1.82
32.	8.35	11.39	15.19	14.	5.18	5.32	12.60
33.	27.26	27.29	27.91	15.	36.00	36.29	38.3
34.	5.24	5.36	12.93	16.	21.89	22.68	28.35
35.	13.97	7.17	1.21	17.	5.57	6.97	6.70
36.	78.05	80.37	85.52	18.	11.83	6.13	1.60

Table II Range of ΔE obtained from the three instruments in four colour difference units.

ΔE		NBS			CIE			FMC-2			Cube-Root		
		Spec	CM	FOC	Spec	CM	FOC	Spec	CM	FOC	Spec	CM	FOC
Paint 1	max	.75	.89	2.86	.89	.92	2.72	2.05	2.11	4.91	.85	.93	2.60
	min	.62	.35	1.11	.80	.34	1.00	1.80	.62	1.56	.75	.34	.99
Paint 2	max	1.54	1.60	3.35	1.60	1.88	3.32	3.98	4.78	5.81	1.59	1.77	2.98
	min	.97	.44	1.47	.98	.66	1.60	2.35	1.53	2.89	.98	.57	1.37
Paint 3	max	2.03	1.89	2.98	1.60	1.04	1.57	3.14	2.88	3.68	1.78	1.69	2.38
	min	1.11	.18	.64	.58	.21	.51	1.60	.33	1.48	.97	.19	.80
Paint 4	max	1.33	1.03	2.15	1.38	1.21	1.88	3.07	2.69	4.05	1.40	1.17	2.01
	min	.26	.36	.53	.26	.37	.56	.52	.75	1.38	.27	.38	.55
Paint 5	max	1.28	1.42	1.50	1.32	1.44	1.48	2.14	2.29	2.33	1.17	1.28	1.32
	min	.26	.18	.43	.26	.19	.67	.38	.41	1.27	.22	.18	.48
Paint 6	max	.87	.92	5.09	.80	.86	5.13	1.69	1.75	8.83	.76	.83	4.53
	min	.29	.54	.84	.27	.57	1.00	.41	1.27	2.10	.25	.55	.88
Paint 7	max	2.30	.88	4.88	2.23	.66	3.42	5.10	2.07	7.58	2.28	.88	4.16
	min	.45	.33	.56	.23	.22	.34	.66	.68	.82	.37	.30	.46
Paint 8	max	1.42	3.33	4.23	1.96	3.47	3.21	5.14	8.07	10.42	1.41	3.28	3.80
	min	.58	2.09	3.20	.64	1.58	2.20	.79	4.90	7.38	.58	1.98	3.03
Paint 9	max	2.11	4.54	3.46	3.75	8.11	5.98	4.43	6.40	4.98	2.22	4.89	3.70
	min	1.36	3.37	1.76	1.98	5.42	3.19	2.62	4.46	2.45	1.40	3.52	1.86
Paint 10	max	1.59	4.76	1.22	1.35	2.63	.82	3.07	13.39	1.96	1.44	5.24	1.04
	min	.19	1.18	.41	.11	.64	.26	.56	2.92	.83	.22	1.17	.38
Paint 11	max	.82	.87	.98	.91	1.12	1.13	1.70	2.11	2.57	.78	.93	1.00
	min	.38	.24	.49	.36	.31	.69	.67	.65	1.48	.32	.27	.56
Paint 12	max	.37	.74	.98	.45	.78	.95	1.11	1.69	1.78	.43	.71	.86
	min	.25	.51	.55	.26	.56	.56	.57	1.16	1.36	.25	.51	.56
Paint 13	max	1.49	3.71	1.20	1.59	4.70	2.20	2.46	5.87	1.73	1.42	3.54	1.18
	min	.27	.82	.60	.20	1.18	.77	.29	2.00	.70	.24	.82	.58
Paint 14	max	2.93	1.20	2.72	2.92	1.05	2.01	8.63	3.29	3.70	3.36	1.29	2.41
	min	.38	.55	1.55	.45	.52	1.20	.65	1.10	2.33	.41	.59	1.49
Paint 15	max	.90	.67	1.07	1.00	.93	1.01	2.41	1.96	2.35	.95	.74	.91
	min	.39	.44	.55	.40	.54	.73	1.02	1.11	1.44	.40	.47	.60
Paint 16	max	.64	.74	2.21	.63	.81	2.14	1.39	1.92	3.75	.62	.76	1.91
	min	.17	.37	.73	.16	.43	.70	.34	.90	1.31	.16	.37	.66

Table II Continued

ΔE	Spec	CM	FOC	Spec	CM	FOC	Spec	CM	FOC	Spec	CM	FOC
	NBS			CIE			FMC-2			Cube Root		
Paint 17 max	.86	.76	5.48	.67	.69	4.95	1.44	1.41	8.32	.75	.71	4.96
min	.56	.40	.28	.48	.40	.23	1.11	.97	.47	.53	.45	.26
Paint 18 max	1.40	4.13	2.10	2.38	8.49	3.01	2.46	7.14	2.57	1.35	4.57	2.18
min	.94	1.34	.69	1.38	.86	1.11	1.00	2.25	.99	.98	1.23	.74
Paint 19 max	1.29	3.33	2.61	1.15	3.30	2.27	1.95	5.34	3.62	1.15	2.82	2.26
min	.54	.80	1.02	.63	.71	1.27	1.27	1.34	2.48	.57	.75	1.10
Paint 20 max	1.12	1.15	3.42	1.27	1.45	3.33	3.43	3.17	4.93	1.28	1.35	2.82
min	.69	.38	1.06	.71	.52	1.05	2.02	1.27	1.70	.74	.51	.91
Paint 21 max	.45	.60	.45	.47	.70	.49	1.18	1.80	1.18	.43	.64	.47
min	.19	.22	.20	.21	.33	.20	.53	.66	.51	.21	.26	.20
Paint 22 max	1.45	3.35	2.21	1.17	2.67	1.99	3.71	11.43	4.49	1.74	4.91	3.50
min	.46	1.03	.75	.36	.62	.41	1.09	1.70	1.90	.58	1.03	1.22
Paint 23 max	1.96	1.13	1.68	2.89	1.26	2.45	6.20	2.30	5.35	2.36	1.04	2.06
min	.43	.26	.51	.52	.26	.47	.98	.59	.84	.41	.26	.47
Paint 24 max	1.22	1.88	2.16	1.52	2.45	1.50	4.28	9.60	11.60	1.44	2.76	3.40
min	.51	1.08	.58	.43	.60	.32	1.83	4.59	2.24	.55	1.50	.66
Paint 25 max	.48	2.22	1.17	.53	1.90	1.10	1.23	3.61	2.42	.49	2.20	1.13
min	.17	1.28	.30	.18	1.23	.28	.47	2.27	.57	.18	1.27	.28
Paint 26 max	2.34	4.19	1.47	3.94	2.88	1.54	2.46	5.10	2.69	2.53	4.19	1.39
min	.67	1.24	.19	1.20	1.31	.10	.65	2.55	.36	.77	1.32	.18
Paint 27 max	.75	.55	3.12	.67	.52	4.60	1.54	1.00	8.76	.68	.48	3.17
min	.49	.34	2.14	.47	.30	3.05	1.17	.63	5.74	.47	.31	2.11
Paint 28 max	.42	.84	.42	.48	.99	.53	.94	1.88	1.12	.50	.99	.59
min	.39	.53	.22	.39	.57	.23	.58	.93	.70	.38	.51	.29
Paint 29 max	3.71	1.14	3.50	2.90	.91	2.11	8.19	3.50	4.60	3.37	1.33	2.79
min	2.78	.29	3.27	2.14	.26	1.99	7.23	.86	4.38	2.74	.31	2.62
Paint 30 max	.64	.55	.88	.63	.55	.86	1.60	1.05	1.63	.63	.51	.76
min	.47	.11	.56	.53	.12	.69	1.24	.29	1.23	.48	.11	.52
Paint 31 max	3.78	2.56	.98	2.38	1.65	.88	8.24	6.04	1.71	3.39	2.54	.92
min	2.53	1.86	.80	1.85	1.43	.61	6.28	4.58	1.30	2.46	2.09	.74
Paint 32 max	1.16	.49	.33	1.14	.52	.36	2.12	1.24	.77	1.14	.61	.40
min	.84	.29	.10	.95	.37	.09	1.90	.82	.14	.90	.42	.09
Paint 33 max	.87	.71	1.42	.92	.97	1.32	1.85	2.16	2.26	.88	.77	1.27
min	.73	.46	.78	.74	.72	.71	1.62	1.52	1.15	.73	.57	.68

Table II Continued

ΔE	NBS			CIE			FMC-2			Cube Root		
	Spec	CM	FOC	Spec	CM	FOC	Spec	CM	FOC	Spec	CM	FOC
Paint 34 max	1.55	1.11	1.04	1.28	1.00	.92	3.68	3.02	1.50	1.56	1.33	.92
min	1.07	.84	.43	1.13	.77	.43	2.60	1.94	1.07	1.12	.97	.48
Paint 35 max	1.96	5.30	1.14	2.26	6.72	1.71	4.45	4.94	1.21	2.12	5.38	1.07
min	1.75	4.33	.77	1.02	5.05	.62	2.15	4.40	.78	1.75	4.42	.87
Paint 36 max	.94	.53	.69	.95	.52	.98	2.27	1.23	2.07	.92	.51	.78
min	.37	.35	.55	.36	.36	.76	.83	.57	1.47	.35	.30	.62

Spec = Spectronic

CM = Colormaster

FOC = Fibre Optics Colorimeter

From these results it is evident that no matter which colour system we use, the acceptable tolerance, expressed as a number of colour difference units, will vary from one part of colour space to another.

These results demonstrate how great the variations between results from different instruments can be. Small variations would be expected since, as has been mentioned above, the instrument responses do not correspond exactly to the CIE standard observer.

The readings from tristimulus colorimeters are known to be of doubtful value for absolute colorimetry, but they are generally believed to give meaningful results when measuring colour differences between nearly non-metameric matches, and results from differential colorimeters in such cases are expected to be in reasonable agreement. This would be the case if the order of colour difference was much greater than the sensitivity of the instrument. However, with industrial finish in paint, the tolerances are very strict and the colour difference between standard and sample is within the sensitivity of most instruments. The tolerances in the paint industry are stated by Tilleard (1964) to be 1–2 NBS units for decorative finish, but better than 1 NBS unit for industrial finish. Thus the large differences in the readings of the two differential colorimeters employed in this study may be attributed to the fact that the colour difference being measured is smaller than the sensitivity of one or both of the instruments.

From the above results, it does not seem possible to specify tolerances for all colours of industrial finish by a single figure. It has therefore been decided to recommend the following practice.

For each colour of industrial finish, or for a range of colours, (e.g. dark reds, light blues), a figure or figures based on a specific instrument can be used

to define the tolerance, provided the instrument is suitable, The tolerance should be expressed in terms of ΔX , ΔY , ΔZ , or in CIE units of colour difference.

The instrument used must satisfy certain specifications. It must give repeatable readings over a long period of time for a permanent standard. (The Fibre Optics Colorimeter has a high degree of reproducibility, due to the built-in calibration based on the direct response of the instrument as a whole through good quality yellow and blue glass filters.) Also the instrument must have a sensitivity greater than that of the human eye under favourable conditions.

If a firm employs more than one colorimeter (for example, there might be a colorimeter in each factory), the instruments should be as nearly identical as possible in their responses. Similarly, to avoid disputes between customer and manufacturer, it is recommended that the instruments used by both should have almost identical responses.

CONCLUSIONS

It is therefore recommended that specification of tolerance for one colour, or range of colours, be based on instrumental measurements, provided that the instrument is more sensitive than the human eye, and that its readings are repeatable. The tolerance may be expressed either in terms of ΔX , ΔY , ΔZ , or in CIE units of colour difference.

ACKNOWLEDGEMENTS

The authors wish to thank the Council and Director of the Paint Research Association, Teddington, England, for permission to publish. Thanks are also due to the member firms who kindly provided the panels for this study.

REFERENCES

- Billmeyer Jr., F. W., (1965), Precision of colour measurement with the G. E. spectrophotometer. *J. Opt. Soc. Amer.* 55, 707-717
- Bond, M. E., and Nickerson, D. (1942), Color-order systems, Munsell and Ostwald. *J. Opt. Soc. Amer.* 32, 709-719
- C.I.E. Proc. CIE (1955), National Secretariat Committee USA on Colorimetry, Sect. 1.3.1
- Friele L. F. C. (1959), Measurement and grading of whiteness of optical bleached materials. *Farbe* 8, 171-186
- Hunter, R. S. (1942), Photoelectric tristimulus colorimetry with three filters. *J. Opt. Soc. Amer.* 32, 509-538
- Ishak, I. G. H., (1970), Fibre Optics Photoelectric colorimeter. *Opt. Acta* 17, 725-732
- Ishak, I. G. H., Bouma, H., and Van Brussel, H. J. J. (1970), Subjective estimates of colour attributes for surface colours. *Vision Res.* 10, 489-500
- Judd, D. B. (1939), Specification of color tolerances at the National Bureau of Standards. *Amer. J. Psychol.* 52, 418-428
- MacAdam, D. L. (1942). Visual sensitivities to colour differences in daylight. *J. Opt. Soc. Amer.* 32, 247-274

- MacAdam, D. L. (1943), Specification of small chromaticity differences. *J. Opt. Soc. Amer.* 33, 18-26
- McLaren, K. (1969), Scaling factors in color difference formulas. *Color Eng.* 7 (6) 38-44
- Richter, M. (1955), The official German standard Color Chart. *J. Opt. Soc. Amer.* 45, 223-226
- Robertson, A. R., and Wright, W. D. (1965), International comparison of working standards for colorimetry. *J. Opt. Soc. Amer.* 55, 694-706
- Thurner, K., and Walther, V. (1969), Untersuchungen zur Korrelation von Farbabstandsbewertungen auf visuellem Wege und über Farbdifferenzformeln. *Proc. 1st AIC congress 'Color 69', Stockholm*, 671-687; *Farbe* 18, 191-206
- Tilleard, D. L. (1964), Evaluation of small colour differences and tolerances in colour matching. *Rev. of current lit. paint and allied ind.* 37, 143
- Wright, W. D. (1941), The sensitivity of the eye to small colour differences. *Proc. Phys. Soc.* 53, 93-112
- Wyszecki, G. (1963), Proposal for a new color-difference formula. *J. Opt. Soc. Amer.* 53, 1318-1319

DISCUSSION

Jaekel: Three comments, on sample selection, colorimeter variations and tolerance-limit variations.

(i) If it had been possible, it would have been useful to measure unacceptable samples, to see whether there were useful ΔE regions that did not overlap, that differed, for passed and failed samples.

(ii) With 45 textile samples in various colours ΔE 's (in ACV units) from the respective standards from Colormaster V and Harrison 70 readings fell within 95% confidence limits of ± 0.35 from the theoretical 45° line. Fairly good for Colorimeter Variation, therefore.

(iii) We are not surprised you to find different tolerances for different colours on one formula. First, we should regard formulae as perfect, and test them as if uniform ΔE tolerance limits would be yielded for all colours by one formula. Then, having selected the more or most promising one or two, we must recognize the imperfections of formulae, and do further work to set appropriate, different, tolerance limits for different colours. Because a formula is imperfect and not truly uniform, this does not mean it cannot be very useful, when employed intelligently.

Ishak: (i) Unfortunately this was impossible because the firms concerned did not keep rejected samples in their files.

(ii) Perhaps 45 samples is rather a limited number from which to draw conclusions.

(iii) Each formula has its uses in a particular field, but use of so many different formulae causes confusion.

McLaren: Whilst I would entirely agree that no formula has yet emerged as the best, I do not know of any evidence to show that CIE 64 is even the best in any series. I therefore find it surprising that Dr. Ishak should recommend this formula in preference to any other.

Ishak: We did not claim that the CIE 64 formulae was superior to any other. Indeed, as has emerged from discussions at this symposium, there is no formula sufficiently good to be recommended for general use. Until such a superior formula is found, in order to avoid confusion, the CIE recommendation could be followed.

Jaekel: I have evidence that CIE 1964 is inferior to other formulae for acceptability decisions.

It is not sufficient to have a good instrument: using it with a bad colour-difference formula, we have shown, can produce worse decisions—more errors—than obtained visually.

Ishak: Each formula has been demonstrated to show superiority over others using a range of colours. An ideal formula has not been found.

We believe that if an instrument is found which satisfies the following requirement:

- a) more sensitive than the human eye
- b) high degree of repeatability

it can be used to make pass/fail decisions.

Kowaliski: The relationship between acceptability and perceptibility is specific to each particular industry.

For this reason it might be useful to check *colour difference formulas* rather in relation to perceptibility and to leave the very involved problems of the related acceptability (and quality control) problems to discussions by specific groups.

Ishak: The usefulness of colour difference formulae for specifying colour tolerances depends on the applications in industry. In industrial finish in paint, perceptibility is synonymous with acceptability, and for this application colour difference formulae may be useful. This may not be true in other cases.

Lozano: When measuring colour with instruments you get considerable variability in the results. The precision of an instrument should be carefully measured when you try to measure colour differences because you do not measure the differences directly but you take one measurement after the other, so the errors are the sum of both. Are we not masking the differences measured with the errors of the measurement? Specially in this case where the differences are so low. There is evidence in the measurements done with the Fibre Optics Colorimeter that its values sometimes get larger than the others. Perhaps that could be explained if the error of this instrument is the largest one?

Ishak: The figures presented here are the averages of at least three measurements. The precision of the Fibre Optics Colorimeter is very high, as is its repeatability, as tested during the present work, and on other samples.

Saltzman: Unlike everyone else in this room I don't know which color difference formula is the best, but I *do* know that you can use color difference formulas (pick your own). I go back to the paper by Nickerson and Stulz (1944)

—and I don't quote it *verbatim*—but it was concluded that there were several usable formulas. The one that you would choose would depend, in part, on the form in which you had your data. While we don't have the perfect formula for color difference, based either on perceptibility or acceptability, at least tens of millions of dollars (I might even make it higher than that) has been paid out on the basis of instrument measurements. Whether these measurements are a mirage I can not say but it's the only thing we have got; it may be a crooked wheel but it's the only game in town.

Color difference formulas do work but they only work if you use them with some intelligence. This is the main thing; if you give some idiot an idiot box (an instrument of unknown quality) you get idiot results.

But if, however, you do know what you are doing as do most of us in the colorant making and consuming industries—you can get useful results by applying color difference calculations to instrumental measurements.

Let us go back a bit. We have been for example dyeing cloth for quite a while and have achieved successful color control with and without instruments. Instruments can, however, help us to do the job better, faster and cheaper. You can and we do use the instruments: If you are making measurements—and I don't care if you are using a ruler, a micrometer or a laser beam interferometer—you first calibrate your equipment and make certain it is working to design-specifications and is sensitive and reliable enough to make the kind of measurements you need (or think you need!). After you are certain of the performance of your instrument you use it to gather data.

In my opinion the greatest problem before us, which limits our work in color difference measurements, is still at the present moment the problem of the sample. We can make measurements with many different kinds of instruments with a precision and reliability that far exceeds our ability to get material which truly represents the sample. For example, in continuous dyeing operations no one will take a statistically reliable sample along the material because it will ruin the piece of cloth and make it unsaleable: How many people take duplicate or replicate measurements? In paint, for example, what is the replicability of our panel preparation even from a uniform liquid sample which truly represents the kettle of paint? In the series described by Dr. Ishak the size of the color differences reported are, if only one sample was prepared and measured, of the same order of magnitude that Miss Ruth Johnston described in her paper "Pitfalls in color specification" (1963). Nevertheless, in spite of the foregoing Pittsburg Plate Glass (PPG Industries), Sherwin Williams, Celanese—to name but three—, successfully monitor and control formulation and production day to day by numerical specification to very close tolerances using instrumental measurement and mathematical treatment of the data.

We, in our own company, sell and get acceptance by the customer on the basis of knowing what we are doing. It is difficult to get to this point, but we

do it and we do it every day. It is not a question in our case and in the case of others of having bad instruments or bad samples; we work all this through. When we do get a difference that falls outside of the tolerance limits accepted by our customer and ourselves we know that it is a real difference. If we were to ship such material we know it will be rejected. We also know that if you fix on a number, by agreement, with which every one is satisfied the day will come when the customer's warehouse is just bulging with both raw material (pigment or dye) and colored textiles and he wants no more of it and so you get a rejection. This is a problem in applied psychology—you send your star salesman or director of sales to discuss this with the customer. But, it will save you the trouble of “going down the garden path” looking for a non existent variation in color! Now lest you think that I am speaking of material made to loose tolerances I should like to state that we sell, among other things, organics pigment dispersions for use in spun dyed textiles. These materials cost up to \$ 20/lb so that a single mix of material (one shipping lot) may represent \$ 50,000-60,000. I know that is not worth as much as it used to be, but it's still a lot of money.

We are shipping this to tolerances of an order of 1.5 to 2 MacAdam units in chromaticity and slightly better than that in lightness. This procedure is reviewed regularly and, as with many specifications the limits have been tightened from the original requirements. So these things do work.

But even with perfect tools you need intelligent people who know something about color and something about your business and then you can use even today's data. Even with perfect formulas no one will buy spectrophotometric curves or color differences.

Hopefully color difference formulas of the future will give us numbers which relate to acceptability but, when business conditions are terrible you will get rejection, when business needs certain colors you can empty a warehouse of non-specification material. If the fashion changes and they need violet, they'll buy every pound of violet dye we have in the place. But this is another problem; let us not confuse this acceptance with meeting specifications. The danger in getting something like that violet accepted, is that somebody will make a note that this was 27 units off and accepted and therefore this company will accept 27 units in the future.

If color difference formulas were no good, people could not do the amount of computer color matching, or computer color correction, and they do lots and lots of it, for money. So let's keep our perspective. improve what we have, but also understand that we have some useful although imperfect tools that can be used if we do it with intelligence.

REFERENCES

- Johnston, R., (1963), Pitfalls in color specification. Official digest 35, 259
Nickerson, D., and Stultz, K. F. (1944), Color tolerance specification. J. Opt. Soc. Amer.
34, 550-570

21 COLOUR DISCRIMINATION STUDIES IN CERAMIC WALL-TILES

F. MALKIN AND A. DINSDALE

*British Ceramic Research Association
Stoke-on-Trent, ST4 7LQ, England*

The purpose of the studies described in this paper was to determine the colour tolerances which should be applied in the production of glazed wall-tiles, and to find the most suitable uniform colour space in which to express both these tolerances and the differences measured during production control.

COLOUR TOLERANCE EXPERIMENT

This experiment was carried out with tiles of five colours, the 1931 CIE co-ordinates of which are given in Table I.

Table I. The 1931 CIE co-ordinates of the five colours studied.

1931 CIE	White	Grey	Pink	Blue	Yellow
<i>Y</i>	83	44	52	31	73
<i>x</i>	0.316	0.318	0.359	0.269	0.355
<i>y</i>	0.322	0.325	0.327	0.287	0.377

Some 600 tiles of each colour were measured on a Model V Colormaster, taking the mean of two readings obtained using the normal aperture and Illuminant 'C', and the results expressed in the modified Adams Uniform Colour Space (Glasser and Troy, 1952). The standard deviation of repeat of the Adams co-ordinates was <0.05 unit for the white, grey and yellow tiles and <0.1 unit for the pink and blue tiles.

Thus, for each colour, we were able to select five or six panels of 48 tiles containing known amounts of colour variation, ranging from a sensibly uniform panel to one so varied that most observers would be expected to reject

it. Care was taken to randomize the colour variation within each panel. The panels, which subtended an angle-of-view of approximately 15° - 20° were mounted on vertical supports which were painted a neutral grey of similar luminance to the tiles. The lighting was natural daylight, mainly from overhead skylights. The 700 observers were asked to indicate any panels which were not acceptable from the point of view of colour variation.

Originally, the percentage rejection votes received by each panel were plotted against the maximum colour difference occurring between adjacent tiles in that panel (Malkin and Dinsdale, 1967). This was unsatisfactory since the colour difference taken for each panel only depended on the measurements on two tiles in the panel, whereas the observers appraised the panel as a whole. Accordingly, a computer program was written to calculate all 82 colour differences and their standard deviation (σ). The total colour difference (ΔE) in each panel was then derived using the formula

$$\Delta E = \text{Mean} + 2.4\sigma$$

and these values of ΔE were compared with the rejection voting. (For a normal distribution of the 82 colour differences, 2.4σ is equal to one half the mean range of difference).

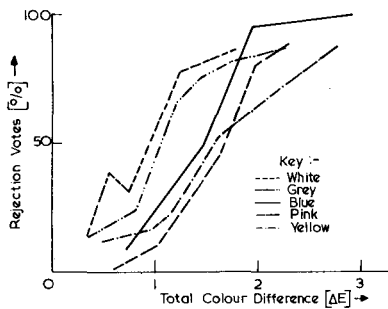


Fig. 1. Tolerance curves, Adams formula

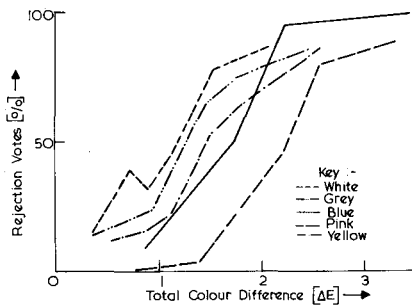


Fig. 2. Tolerance curves, CIE 1964 formula.

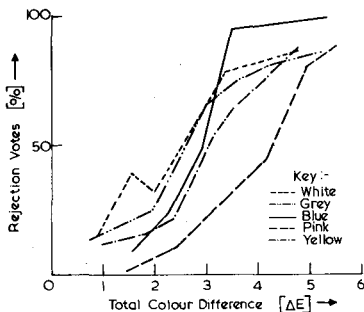


Fig. 3. Tolerance curves, S. & G. formula.

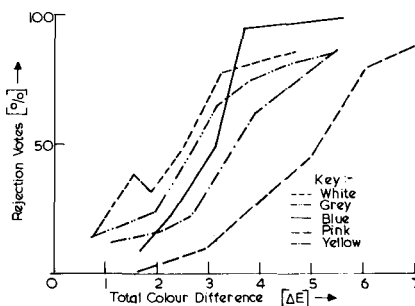


Fig. 4. Tolerance curves, FMC-2 formula.

Fig. 1 shows the correlation graph for the five colours with total colour differences calculated using the modified Adams formula. The tolerance for white and grey is seen to be significantly less than that for the three pastel colours. The irregularity in the voting for the second white panel was found to be spurious. At certain times of the day, this panel acquired an unpleasant yellow cast due to sunlight reflected from adjacent building supports causing some observers to reject it for this reason, and not because of its colour variation.

It seems that the Adams formula under-rates colour difference for neutrals compared with those for pastel colours. The colour differences were therefore re-calculated using four other formulae to see if they gave better agreement with the observer's voting. The four formulae tried were: the CIE 1964 (Wyszecki, 1963); a computer version of the Simon and Goodwin Charts (Simon and Goodwin, 1958); the Reilly cube-root formula (Glasser, Reilly *et al.*, 1958; and Wyszecki, 1968); and the Friele-MacAdam-Chickering II formula (Wyszecki, 1968). Tolerance graphs were plotted in exactly the same manner as for the Adams formula.

The graph for the CIE 1964 formula is shown in Fig. 2. The tolerance curves for blue, yellow, white and grey tiles are closer together on this plot, but the curve for pink tiles is separated from them. Thus the CIE 1964 colour difference formula offers no advantage compared with the modified Adams formula.

Fig. 3 shows the tolerance graph obtained using the Simon and Goodwin method. This time there is a marked improvement in the spacing of the curves for four of the colours, the exception being the pink curve. But for this, the Simon and Goodwin method would show a significant improvement over the modified Adams formula.

The Reilly cube-root formula yielded tolerance curves which were very similar to those for the Adams formula (Fig. 1). This formula, therefore, does not appear to offer any advantage over the Adams one.

The FMC-2 formula produced tolerance curves shown in Fig. 4, which had the same pattern as those derived from the CIE 1964 formula (Fig. 2). The FMC-2 formula exaggerated the tolerance for pink more than any other formula examined.

Summarising, we can say that the experiment gave a useful indication of customer tolerance for the colours studied, but that none of the colour difference formulae examined gave very good agreement with the observer's assessment. The extent of this disagreement is seen in Table II, where the colour differences which received a 50% vote are compared.

There is no previous report of the discrepancy in tolerance between pastel colours and neutrals which is evident in these results. However, examination of the work of Davidson and Friede (1953) on the tolerance in rug wools reveals some evidence which tends to confirm this discrepancy. The 19 colours which they studied can be divided into four groups according to their Munsell Chroma.

Table II. Colour Differences which received a 50% rejection vote.

Colour Difference Formula					
Colour	Adams	Cube-root	CIE 1964	S & G	FMC-2
White	0.9 ₅	0.9 ₅	1.2	2.6	2.5
Grey	1.0 ₅	1.1	1.2 ₅	2.6	2.7
Pink	1.7	1.8	2.2	4.3	5.1
Blue	1.5	1.5	1.7	2.9	3.1
Yellow	1.6	1.5 ₅	1.5	3.1	3.5

Fig. 5 shows the correlation graphs plotted for each group for colour differences expressed in the Adams formula. In spite of the experimental scatter of each graph, it is clear that the general slope of the Chroma <2 graph is steeper than the others, indicating a reduced tolerance for the neutrals and very weak colours similar to that observed for wall-tiles. This analysis of the Davidson and Friede data has been confirmed by the multiple linear regression carried out by McLaren (1971).

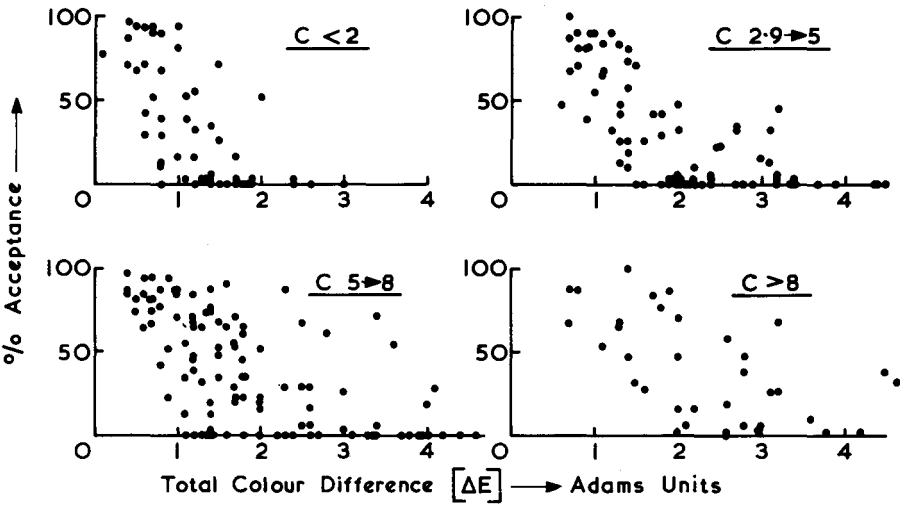


Fig. 5. Analysis of the Davidson and Friede results, Adams formula.

PINK COLOUR-DIFFERENCE SCALE EXPERIMENT

The anomaly in the tolerance for pink tiles prompted further investigation. The first step was to calculate the relative weighting which each formula gave to different types of colour difference in the pink region. The Adams Space was taken as reference, and unit differences in the directions, *L*, *a*, *b*, hue and

saturation about the pink tile colour in Adams Space were re-calculated using the other four formulae tested. The results, given in Table III are most inter-

Table III. ΔE values for colour differences in pink tiles.

<i>Colour difference</i>	<i>Adams</i>	<i>Cube-root</i>	<i>CIE 1964</i>	<i>S & G</i>	<i>FMC-2</i>
Adams 'L'	1.0	1.0	1.0	2.6	2.5
Adams 'a'	1.0	1.1	1.6	2.6	3.1
Adams 'b'	1.0	1.0	1.2	1.9	2.2
Hue	1.0	0.85	1.0	1.4	1.5
Saturation	1.0	1.15	1.7	2.7	3.3

esting. In particular they reveal the very different weightings which are given to hue and saturation differences in pink by the various formulae. It was decided to carry out an experiment to find out which of these weightings agreed with visual experience.

Sets of pink tiles were selected to make a colour difference scale against which other colour differences could be compared. The scale contained colour differences ranging from 0.4 to 1.7 Adams units in seven steps, all in the same direction in colour space, namely that of the Adams 'b' axis. In each set a tile having the required colour difference was displayed between two standard tiles. This arrangement was found to give much more reliable judgements of small colour differences than could be obtained by using only one standard, as it largely overcame the bias often observed when only one standard was used. This bias, which is believed to be due to retinal asymmetry, caused observers to assess the right-hand tile of an identical pair as being, say, paler than the other, and to persist with this assessment even when the tiles were interchanged.* Twenty four other sets of pink tiles, exhibiting different amounts and types of colour difference, were assembled for comparison with the scale. All tiles used were checked for uniformity of colour, to ensure that the colour differences measured at the edges of adjacent tiles in a set agreed with those measured between the centres of the tiles. (This criterion was relaxed slightly for five of the comparison sets, with the interesting results noted below).

The colour difference scale was displayed on a vertical board which was painted to be very similar in colour to the tiles, the remainder of the field of view was neutral grey. Two banks of artificial daylight fluorescent lamps, British Standards 950 part I (1967), illuminated the scale at 45°, with an intensity of illumination of 75 lumens per square foot. Viewing was normal to the tile surface. Each set of tiles subtended an angle of view of 9° to the observer,

* Dr. F. J. J. Clarke has pointed out that this retinal bias effect is well known in the field of photometry. The instructions for use of the Lummer-Brodhun Photometer specify that the role of sample and reference fields should be interchanged to overcome this bias effect.

and the pink board subtended an angle of 25° . The colour vision of the thirteen observers was checked using the Ishihara Charts and the Farnsworth-Munsell Test.

To facilitate judgement, the comparison sets were placed between sets on the scale, as requested by the observer. One comparison set was repeated at intervals during the experiment as a control. Each observer estimated the magnitude of the total colour difference in the 24 comparison sets, expressing his results to $\frac{1}{2}$ step on the scale 1–7. (The few assessments greater than 7 were rated at $7\frac{1}{2}$). Results on this scale were converted to equivalent ΔE values, Adams units, by reference to the known values of ΔE for each set on the scale. For each comparison set we could then calculate the ratio of its measured colour difference, $(\Delta E)_m$, to the visually equivalent difference, $(\Delta E)_v$, assessed by the observer. This ratio is a measure of the magnitude of colour difference, which is judged to be visually equivalent to 1 unit of difference on the scale i.e. to 1 unit in the direction of the 'b' axis of the Adams space. By expressing the results in terms of this ratio, all judgements on one type of colour difference could be grouped together, irrespective of the magnitude of the individual differences.

The mean value of this ratio was determined for each comparison set. In general there was good agreement between the mean values obtained on sets which exhibited the same type of colour difference. The exceptions were the five sets, noted above, containing tiles which were not uniform in colour. For example, eight judgements on sets with uniform tiles, having the same type of colour difference as the scale, yielded mean ratios from 0.97 to 1.12; while a similar set containing non-uniform tiles gave a mean ratio of 0.67 if centre differences were used in the calculation, but a value of 0.95 if the maximum edge difference was used. These results provide strong evidence that judgement of colour difference is based on the difference occurring at adjacent edges of samples, and not on the difference between the centres of the samples. This is in accord with the demonstration of graded samples by Balinkin (1965), and with the findings of Land (1971).^{*} In view of this, the maximum edge difference was used when calculating the ratio for the five sets containing non-uniform tiles.

The results of the observer's judgements on separate lightness and chromaticity differences are summarised in Table IV. For each type of colour difference, the range of the mean values of the ratio determined for individual sets is given, together with the overall mean for that type of colour difference. The distributions of the individual observer's assessments on the sets are shown in the histograms of Fig. 6A and 6B. The ratios obtained for 'L', 'a', 'b' and saturation differences are all close to 1, showing good agreement with the spac-

^{*} See also the yellow tile demonstration described in the Appendix.

ing of the Adams formula. However the mean value of 0.75 obtained for the hue difference set is significantly different (at the 1% confidence level) from 1. This is in accord with the general comment in industry, and with the findings of Kuehni (1970), who analysed the Thurner and Walther (1969) data to show that tolerances for hue differences were much smaller than those for saturation differences.

Table IV. Judgements on separate lightness and chromaticity differences for pink tiles.

Colour difference	No. of sets judged	Ratio $(\Delta E)_m/(\Delta E)_u$ for 13 observers	
		Range of mean values	Overall mean
Adams 'L'	6	0.90—1.26	1.08
Adams 'a'	4	0.99—1.07	1.02
Adams 'b'	9	0.90—1.12	1.02
Saturation	3	0.89—1.00	0.95
Hue	1	—	0.75

To assess the performance of the other three formulae on these differences, the ΔE values quoted in Table III can be used to calculate the magnitude of each type of colour difference which should be visually equivalent to 1 (Adams) unit of difference in the direction of the Adams 'b' axis. All three formulae give a value of 0.7 for a difference of saturation, which is smaller than the observed value of 0.95; while they give values of between 1.2 and 1.5 for a hue difference, much greater than the observed value of 0.75. Thus, the CIE 1964, the Simon and Goodwin, and the FMC-2 formulae all give weightings for hue and saturation differences which are in direct contradiction to visual observations on pink tiles, and to the results of the Thurner and Walther experiments.

The remaining five sets which were compared with the scale all exhibited combined lightness and chromaticity difference, in which the proportions of lightness and chromaticity were roughly equal. The histograms of the 13 judgements on these sets, shown in Figure 6B, reveal some rather disturbing tendencies. Table V shows the standard deviation (σ) of the ratios obtained for each set, and their mean value, together with the significance of the departure of this mean value from the unit value which would indicate perfect agreement

Table V. Judgements on combined lightness and chromaticity differences.

Colour difference	Mean Ratio <i>M</i>	σ	Significance of $(M-1) \neq 0$		
			5%	1%	0.1%
—S, —L	1.2	0.3 ₇	—	—	—
—S, +L	0.8	0.3 ₂	Yes	—	—
+S, —L	0.8	0.2 ₄	Yes	—	—
+b, +L	1.5	0.2 ₀	Yes	Yes	Yes
+b, —L	0.7	0.2 ₁	Yes	Yes	Yes

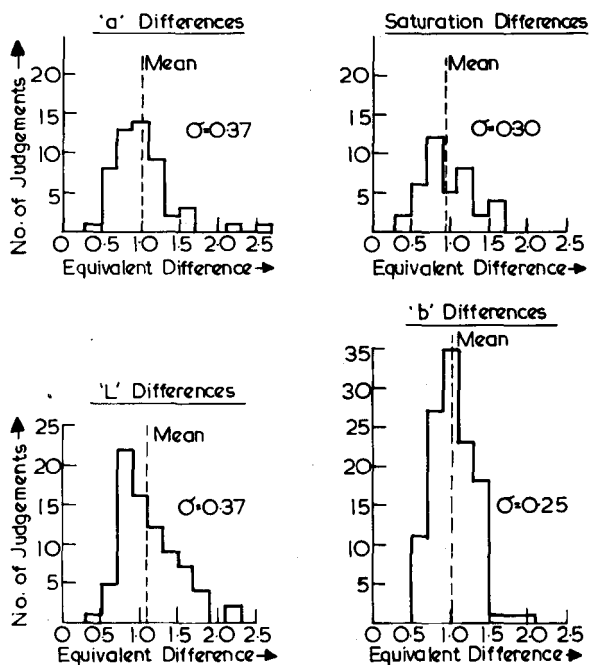


Fig. 6A. Differences judged equivalent to 1 unit, Δb , Adams: pink tiles.

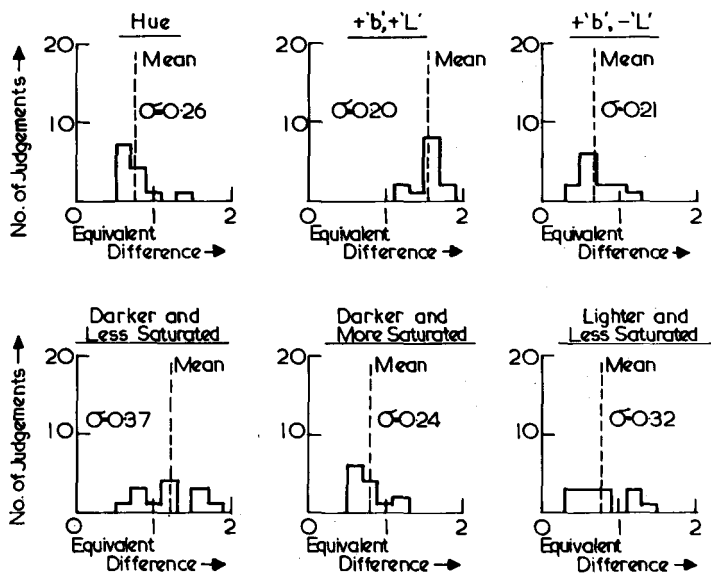


Fig. 6B. Differences judged equivalent to 1 unit, Δb , Adams: pink tiles.

with the Adams colour space. The discrepancies found for the combined lightness and saturation differences are barely significant, but those found for the combined $+ 'b' \pm 'L'$ differences are highly significant, indicating a discrimination ellipsoid which is not symmetrical about the chromaticity plane.

Thus, the modified Adams formula fails to give correct weighting to these combined chromaticity and lightness differences. The other formulae cannot give any better performance in this respect, since in no case does the sign of the lightness component of the difference affect the magnitude of the total colour difference which is calculated by the formula.

Summarising, the Adams and Cube-root formulae give a spacing which is in good agreement with visual observation on separate lightness and chromaticity differences in pink tiles, the only exception being in the case of a hue difference. The other three formulae give much worse spacing. Our observers' assessments of combined lightness and chromaticity differences do not agree with the spacing of any of the five formulae, or with the findings of Brown and MacAdam (1949), whose discrimination ellipsoids were nearly symmetrical about the chromaticity plane.

COLOUR DIFFERENCES IN GREY TILES

In view of the above discrepancy, it was decided to obtain visual assessments on combined lightness and chromaticity differences in a different colour. Three sets of grey tiles were selected. Again each set comprised three tiles, with a 'difference' tile placed between two standards. All the tiles were carefully checked for uniformity of colour. All three sets contained a total colour difference of 1 Adams unit. In set A the centre tile was simply bluer than its standards; in set B the centre tile was bluer and darker than its standards, in approximately equal proportions: and in set C the centre tile was bluer and lighter than its standards, again in equal proportions.

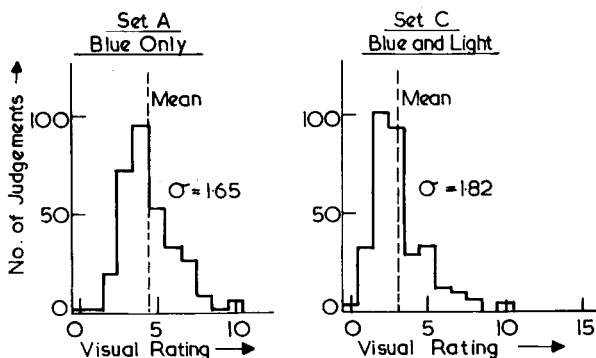


Fig. 7. Visual Ratings of Grey Tile Differences.

The sets were displayed on a matching grey background, otherwise the lighting and viewing conditions were exactly the same as for the pink scale experiment. The 300 observers assessed the magnitude of the total colour difference in sets *A* and *C* by comparison with set *B*, which was given the arbitrary value of 5 units. The histograms of the individual assessments are shown in Fig. 7. Table VI contains the mean rating given to each set and the standard deviation of the judgements, together with the ΔE values calculated on each of the colour difference formulae.

Table VI. Visual ratings and calculated ΔE values for Grey Tiles

Colour difference	ΔE calculated on each formula				Mean Visual Rating	σ
	Adams and Cube-root	CIE 1964	S & G	FMC-2		
Blue only	1.0	1.1	2.4	2.2	4.5	1.65
Blue, Dark	1.0	1.1	2.4	2.3	5	—
Blue, Light	1.0	1.1	2.5	2.5	3.2	1.82

The values given by the various formulae do not agree with the visual ratings, especially for the blue and light set. Thus we find a similar discrepancy for grey tiles with complex colour differences to that observed for pink tiles.

COLOUR DIFFERENCES IN GREEN TILES

Further visual assessment of combined lightness and chromaticity differences have been obtained using green tiles with the approximate CIE co-ordinates $Y = 55$, $x = 0.310$ and $y = 0.360$. Four sets of tiles were judged, each set consisting of a block of four tiles with a pair of standards and a pair of difference tiles arranged in a diagonal pattern. The total colour difference in each set was one Adams unit. The sets contained differences in saturation only ($-S$), lightness only ($+L$), and combinations of saturation and lightness in equal proportions ($-S, +L$) and ($-S, -L$).

The sets were displayed on a background of green tiles with a painted green surround subtending some 40° to the observer. (Each set of four tiles subtended $6^\circ - 7^\circ$). The lighting and viewing conditions were the same as for the pink and grey tile experiments. The ($-S, +L$) set was given the arbitrary value of 5 units, and the 65 observers assessed the magnitude of the colour difference in the other three sets by comparison with it. The histograms of the individual assessments are shown in Fig. 8. Table VII contains this mean rating given to each set and the standard deviation of the judgements, together with the ΔE values calculated on each of the colour difference formulae examined.

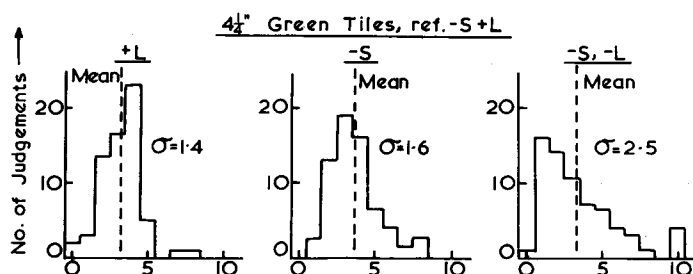


Fig. 8. Visual Ratings of Green Tile Differences.

Table VII. Visual ratings and calculated ΔE values for Green Tiles

Colour difference	Adams	Cube-root	CIE 1964	S & G	FMC-2	Mean visual rating	σ
+L	1.0	1.0	1.0	2.6	2.5	3.3	1.4
-S	1.0	0.85	1.0	1.9	1.7	3.7	1.6
-S, -L	1.0	0.9	1.0	2.0	1.9	3.3	2.5
-S, +L	1.0	0.9	1.0	2.4	2.3	5	—

As with the pink and grey tiles, the calculated ΔE values do not agree with the visual assessments. In all three cases the results suggest that the discrimination ellipsoid is not symmetrical about the chromaticity plane*. There is no previous report of this kind of discrepancy, but there would seem to be some connection with the work of Crawford (1969) who comments p. 305, 'The complete relation between perception of a brightness difference and of a colour difference is far from simple; in some cases there appeared to be degradation in perception of a colour difference in presence of a brightness difference, even though the latter had not been perceived.'

CONCLUSION

The colour tolerance experiment gave useful information on customers' requirements, but showed that the colour difference formulae examined did not give good agreement with visual assessment. Smaller tolerances were found for neutrals than for pastel colours, and a similar tendency can be detected in the Davidson and Friede results. The explanation for this might lie in the smaller tolerances found for hue differences than for saturation differences in the Turner and Walther experiments. For pastel colours chromaticity differences may have the appearance of either hue or saturation differences; but for neutral greys, *all* chromaticity differences will have the appearance of hue differences, with a consequent reduction in customer tolerance.

* See also the green tile experiments described in the Appendix.

Only the Adams and Cube-root formulae gave reasonable agreement with visual assessment of the magnitude of separate lightness and chromaticity differences in pink tiles. The other three formulae gave significantly greater weight to saturation differences than to hue differences, in direct contradiction to visual assessment.

None of the formulae agreed with observations on combined lightness and chromaticity differences in pink, grey or green tiles. This result is at variance with previously published data, and the matter therefore needs further careful investigation.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the assistance of Mr J. F. Birtles and Miss D. L. L. Salt in the preparation and running of these experiments, the cooperation of Mr G. N. Vaughan who wrote the necessary computer programs, and the patient collaboration of their observers. They are also pleased to acknowledge the permission of their Director, Dr N. F. Astbury, C. B. E., to present this paper.

REFERENCES

- Balinkin, I. (1965) Keys to color education. *Farbe*, 14, 107-122
- Brown, W. R. J., and MacAdam, D. L. (1949) Visual sensitivities to combined chromaticity and luminance differences. *J. Opt. Soc. Amer.*, 39, 808-834
- British Standards 950 part I (1967). Artificial Daylight for the assessment of colour. British Standards Institution. London
- Crawford, B. H. (1969) Just perceptible chromaticity shifts. *Proc. 1st AIC Congress "Color 69"*, Stockholm, 302-311
- Davidson, H. R., and Friede, E. (1953) The size of acceptable color differences. *J. Opt. Soc. Amer.*, 43, 581-589
- Glasser, L. G., and Troy, D. J. (1952) A new high sensitivity differential colorimeter. *J. Opt. Soc. Amer.*, 42, 652-660
- Glasser, L. G., McKinney, A. H., Reilly, C. D., and Schnelle, P. D. (1958) Cube-root color co-ordinate system. *J. Opt. Soc. Amer.*, 48, 736-740
- Kuehni, R. (1970) The relationship between acceptability and calculated color differences on textiles. *Color Eng.* 8 (1), 47-53
- Land, E. H., and McCann, J. J. (1971) Lightness and retinex theory. *J. Opt. Soc. Amer.*, 61, 1-11
- Malkin, F., and Dinsdale, A. (1967) Colour tolerance studies with ceramic wall-tiles, in *Colour Measurement in Industry*, The Colour Group (G. B.), 221-233
- McLaren, K. (1971) Multiple linear regression: a new technique for improving colour difference formulae. This symposium.
- Simon, F. T., and Goodwin, W. J. (1958) Rapid graphical computation of small color differences. *Amer. Dyestuff Rep.* 47, (4), 105-112
- Turner, K., and Walther, V. (1969), Untersuchungen zur Korrelation von Farbabstandsbewertungen auf visuellem Wege und über Farbdifferenzformeln. *Proc. 1st AIC congress 'Color 69'*, Stockholm, 671-687; *Farbe* 18, 191-206
- Wyszecki, G. (1963) Proposal for a new color-difference formula. *J. Opt. Soc. Amer.*, 53, 1318-1319

Wyszecki, G. (1968) Recent agreements reached by the colorimetry committee of the CIE. *J. Opt. Soc. Amer.*, 58, 290-292

APPENDIX

By kind permission of our Dutch hosts, the following experiments and demonstrations were presented in conjunction with this paper.

Experiments 1 and 2. Green Tiles

Two simple experiments on colour-difference assessment were presented, arranged on the lines of the green tile experiment described in the paper but using smaller samples cut from tiles. The samples were assembled into sets of four with the cut edges adjacent, but separated by thin spacers introduced to provide a small neutral gap between them. (An attempt to assemble the samples in as close contact as possible, following the generally accepted conditions for assessing colour differences, revealed the disturbing effect shown in the second Demonstration).

The mean results obtained by 28 observers at Driebergen were in very close agreement with those obtained by 65 observers at B. Ceram. R. A. They were similar to those for the larger tiles reported in the paper, except that the visual rating for the $-S$, $+L$ set was reduced when the smaller samples were used. The reason for this discrepancy is not known, but may be related to the difference in the nature of the gaps between the tiles in the two cases.

Demonstration 1. Yellow Tiles

This demonstrated the effect of gradation in colour on the magnitude of the difference perceived between two tiles. Two pairs of yellow tiles were displayed on a grey background, each pair having the same colour difference measured between the centres of the tiles. The tiles in one pair were uniform in colour, so that the difference between the adjacent edges of the tiles was the same as that between their centres. This difference of 2 CIE units was clearly visible.

The second pair of tiles were graded in colour, so that the difference between the adjacent edges of the tiles was very much less than that between their centres. In this case no difference was observed between the tiles until a grey card was placed over the adjacent halves of the tiles.

Thus, the colour difference perceived was that occurring at the adjacent edges of the tiles, and not that measured between their centres.

Demonstration 2. Green Tiles

This demonstrated the effect of the gap between samples on the magnitude of the difference perceived between them. Two green colour difference sets ($\Delta E = 2.3$ CIE units) were displayed horizontally on a grey background. In one set, *A*, the gap between the two halves was kept to a minimum by placing the ground edges in close contact. In the other set, *B*, a small gap was introduced between the two halves. Except for gap size the two sets were identical, being cut from the same uniform tiles.

When the sets were viewed at 45° from a distance of about 1 metre, the gap in set *B* was clearly visible and the colour difference in this set was easily seen. However, the gap in set *A* was not visible under these conditions, and the majority of observers either saw no colour difference, or a very much reduced colour difference, in this set.

Thus the presence or absence of a small gap between two samples considerably influenced the magnitude of the colour difference perceived between them. This poses the question, what is the optimum size of gap for viewing colour differences?

DISCUSSION

Jaekel: I agree that direction of difference, certainly in pinks, and probably in other colours, influences acceptability/ ΔE levels. On reading your paper, we looked at our pink textile patterns which we had divided into those predominantly judged yellower, bluer, stronger, and weaker than standard. Colour differences around 3.5 ACV-units with $k = 50$ corresponded to well over 50% acceptance for the stronger pinks, and to less than 20% acceptance for the other three lots. The presence of these difficulties and complexities in formulae doesn't mean, however, that we cannot employ them usefully.

Malkin: I agree. The important thing, though, is to recognize that these complexities do exist.

Simon: Are not you suggesting that the problem lies in the weighting brightness versus chromaticity? I will discuss this important problem this afternoon. When you have these differences in $+S$ or $-S$ and L you might have a different factor for ΔL in the ΔE computation and get a better result. This does not explain the pink tiles, I am sure. In some of these others, though, where you have $+S$ and $-S$ versus L changes, you can weigh these formulas perhaps differently and get calculations more in agreement with your judgements.

Malkin: I don't think we are working on the same wavelength. The colour differences in one set of tiles is $-S, -L$, in the other set $-S, +L$. In fact, within measurement error, on the Adams system those sets have identical colour differences except for the sign of the lightness component. So, the different weighting of lightness will not explain our results.

Simon: Right.

Boynnton: I would like to comment on the gap effect, which Mr Malkin and I discussed at some length last night at dinner, and end with a question for the audience. In our experiments on the minimally-distinct border, we have observed the same phenomenon. We have precise control of the spatial position of our sectors, since the first step in making a minimally-distinct border setting is to juxtapose the fields as accurately as possible.

With low-contrast pairs, the appearance of colour difference between the

two sectors may be dramatically enhanced by separating the fields slightly, thus producing a very fine black line between them. My question to the audience is: Does anyone know of any previous report of this phenomenon?

MacLaren: In industry we often hear that if we cut one sample in two, observers classify the two halves as being distinct.

A way to test this imaginary difference is to have the observers state which sample is more bright than the other and afterwards present the two samples in random configuration. I wonder whether Prof. Boynton has tried this out with his minimal distinct border apparatus.

Boynton: No.

Malkin: Can I answer Mr MacLaren's question?

We have observed the bias effect which he describes when comparing identical samples, but we have found it to be small, of the order of a just noticeable difference. The colour difference displayed in the "gapeffect" demonstration is much larger than this, being greater than 2 CIE units.

Saltzmann: In the paint industry the moral is, if you run out of paint, never run out in the middle of a flat wall. Get yourself to the edge no matter what, and then it makes no difference. You will not have any trouble on the other wall. Even a thin dividing line decreases the differences.

Malkin: I agree entirely, but apparently the line can be too thin. I would think that complete absence of a dividing line is not the ideal condition for judging differences. You need a very fine line of separation to see the differences at their best.

Hunter: May not the fine dividing line which is found necessary to see very small color differences actually serve the visual function of identifying the locations of the two areas being compared?

Malkin: During discussion on the gap effect with Prof. Boynton, I suggested that the effect is largely psychological and he thinks it goes deeper than that.

Boynton: Did I say it that way?

Malkin: Sorry. My point was that, if we are not aware of a clear dividing line between two samples we tend not to make an assessment of colour difference, rather we see the two samples as one uniformly coloured area. Prof. Boynton thought that this was an over-simplification.

Boynton: Mr Malkin mentioned that I had expressed some ideas about the basis of the gap effect, and that perhaps I would comment on them. This was the essence of my reply:

The contours that we perceive in the visual field, which permit the discrimination of objects, seem to be provided mainly by achromatic differences. Thus the perceivable detail in a black-and-white photograph is little less than in a color reproduction. Color seems to fill in the areas defined by these achromatically-defined contours.

In general, the perceived homogeneity of such areas is greater than it has

any right to be given the physical inhomogeneity typically present, and the drastic non-uniformity of the retinal mosaic. The system somehow seems to average the color-signals derived from within the defined area. In our experiment when the minimally distinct border has been set for small chromatic differences, with precise field juxtaposition, the defined area appears to include both sectors, seen as a semi-circle with only a minor discontinuity in the middle, and a small apparent color difference between the two sectors. If now one of the sectors is moved slightly with respect to the other one, introducing a slight gap, the apparent color difference between the two sections is significantly enhanced.

The gap effect appears to be a non-monotonic function of gap width, since for wider gaps, the greater the gap, the less is the perceived color difference. This is probably related in some way to the lateral extent of horizontal mechanisms in the retina of the eye, which operate most effectively, in detecting differences, over short distances.

Jaeckel: Is it possible that the gap effect is dependent on the colour difference of the presented tiles?

Malkin: The difference between the two tiles in the demonstration is not small. It is more than two CIE units, which would be disastrous in tile production!

Wright: The effect reported by Mr Malkin that a fine dividing line between the samples being compared increases the sensitivity of the eye to the detection of small colour differences is clearly very important. It conflicts, for example, with the practice usually recommended in the design of visual photometers and colorimeters to use a photometer prism with as invisible a dividing line between the two fields as possible. The explanation of the effect must presumably be found in the neural response of the retina and may be related to the role of eye movements in discrimination.

Ditchburn has suggested on the basis of experiments with stabilized retinal images that we see colours essentially through retinal reactions occurring at the edges or boundaries of an area of a colour. If an edge cannot be perceived between two fields, this may for some reason lower the discrimination sensitivity, although if the fine edge is broadened into a band, the discrimination will fall off.

Nickerson: To make sure that your tiles were uniform, you had to measure it, as your subjects could not judge it?

Malkin: They certainly could not, as they could not in the case of the yellow tiles in the first demonstration. So indeed, we had to measure the centre and edges of the tiles to check their uniformity of colour.

22 THE USEFULNESS OF COLOUR-DIFFERENCE FORMULAE FOR FIXING COLOUR TOLERANCES*.

W. SCHULTZE

*Badische Anilin- & Soda-Fabrik AG
Ludwigshafen/Rhein, Germany*

The agreement between the assessments of colour matchers and the values obtained by the use of various colour-difference formulae has been investigated by a number of workers. We have taken the experimental results from seven papers and compared them with the results obtained by using ten (or sometimes more) different formulae, using a new criterion to measure the degree of agreement. The publications considered are those of: Nickerson and Stultz (1944), Davidson and Friede (1953), Coates and Warburton (1968) and Coates, Day and Rigg (1969), Robinson (1969), Thurner and Walther (1969), and Schultze and Gall (1969).

1. METHODS OF TESTING COLOUR-DIFFERENCE FORMULAE

In some of the papers cited the experiments covered several regions of colour space, and it was sought to test the formulae for their usefulness in all regions. Certainly the ideal formula would agree with visual assessments in all regions, but we have come to the conclusion that a formula is to be regarded as useful if it agrees well with visual assessments only after applying a correction factor that may vary from region to region. We have therefore considered each region separately in testing the formulae.

Nickerson and Stultz, Robinson, and Thurner and Walther used a scoring system to quantify the visual assessments (scores 1-5 or 1-6). How should the scores obtained be compared with the ΔE -values calculated from instrumental measurements? The problem is essentially the same as that of comparing the ΔE -values obtained by different formulae.

Several authors have used conventional correlation coefficients, but two

* More extensive data will be published in *Die Farbe*.

examples will serve to show that this method of treatment is not the appropriate one.

Example (a)

Fig. 1 shows six colours whose x, y -coordinates lie on a circle with centre S , the coordinates of a standard colour. The ellipse A with axis-ratio 3:1 is the locus of colours with constant colour difference $\Delta E_A = 1$. The curve B is approximately elliptical (ratio 3:2) and is the locus of colours with constant colour difference $\Delta E_B = 1$. The colour differences of the six colours are 1.0, 1.4, 1.8, 2.2, 2.6, 3.0, in accordance with formula A and 1.0, 1.1, 1.2, 1.3, 1.4, 1.5 in accordance with formula B . These are both arithmetical series, and of course the correlation coefficient is 1. However the existence of perfect correlation between the two series of values tells us nothing about the differences between the two formulae. These differences are in fact approximately those that exist between the Godlove and Munsell-Renotation formulae.

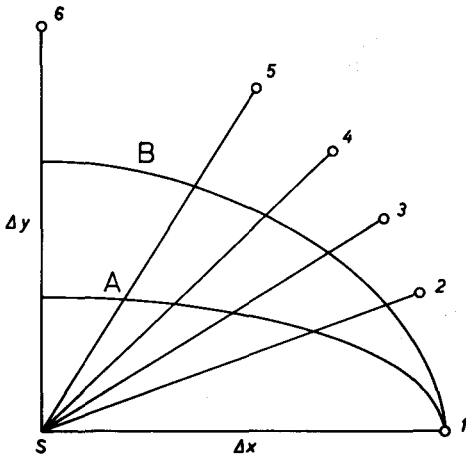


Fig. 1. Illustration of the impropriety of conventional correlation techniques for comparing ΔE formulae.

Example (b)

Suppose that these same six colours are found to have colour differences ΔE_C equal to 1.0, 0.7, 1.4, 1.7, 1.2, and 1.5. The correlation coefficient with respect to series B is only 0.635. The same correlation coefficient would of course be obtained if we doubled or trebled each value ΔE_B and ΔE_C . However, if we pooled the three sets of values ΔE_B and ΔE_C , $2\Delta E_B$ and $2\Delta E_C$, and $3\Delta E_B$ and $3\Delta E_C$ the correlation coefficient would be much higher: 0.894. This result makes no sense in this particular context. The introduction of doubled and trebled values should not affect the relation between the two formulae.

McLaren (1969) suggested a much improved method of comparing two sets of values ΔE_A and ΔE_B . This gives results very similar to our own, but the for-

mula used lacks symmetry, so that different coefficients are obtained when ΔE_A and ΔE_B are exchanged. We have therefore used the following formula in comparing colour-difference formulae (Schultze, 1969 *a* and *b*). The discrepancy between the results obtained by two difference formulae is given by

$$V_{A,B} = \left[\frac{1}{n} \sum_{i=1}^n \frac{(\Delta E_A - F \Delta E_B)^2}{\Delta E_A F \Delta E_B} \right]^{1/2}, \text{ where}$$

$$F = \left[\frac{\sum_{i=1}^n \frac{\Delta E_A}{\Delta E_B}}{\sum_{i=1}^n \frac{\Delta E_B}{\Delta E_A}} \right]^{1/2}$$

$$\text{Eliminating } F \text{ gives } V_{A,B} = \left[\frac{2}{n} \cdot \left(\sum_{i=1}^n \frac{\Delta E_A}{\Delta E_B} \cdot \sum_{i=1}^n \frac{\Delta E_B}{\Delta E_A} \right)^{1/2} - 2 \right]^{1/2}$$

This expression is unaffected by interchanging all the values ΔE_A and ΔE_B . Its minimum value is zero (every ΔE_A proportional to every corresponding ΔE_B). The examples (*a*) and (*b*) considered above give $V_{A,B} = 0.237$ and $V_{B,C} = 0.232$. Since the only terms apart from n are quotients the value of the expression is unaffected by multiplying each pair of values by the same factor or by combining such multiplied pairs with the series.

The quantitative treatment of scores obtained from visual assessments and ΔE -values obtained by using a particular difference formula is in principle the same as that for two series of ΔE -values obtained with two difference formulae. However, there is one important difference between visual scores and ΔE -values. The latter always increase almost linearly as the distance from the reference point increases in a given direction (provided the distance is not too great), although different formulae may give different constants of proportionality. Visual scores on the other hand may increase non-linearly.

Fig. 2 shows possible relationships between visual scores and ΔE -values: the full curve represents a linear relationship, and the pecked curve shows a hypothetical non-linear relationship. Unfortunately none of the publications in which a scoring system is described includes a series of measurements and assessments for colour variation in one direction. Close examination of the data given by Robinson and by Thurner and Walther does however indicate that the visual

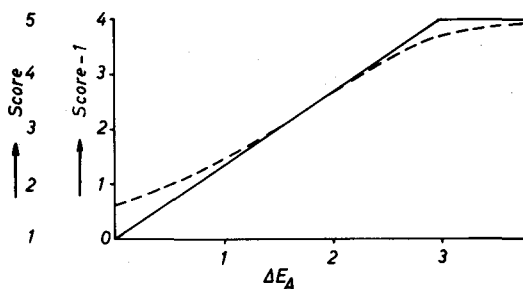


Fig. 2. Hypothetical relations between visual assessment score and calculated colour-difference.

scores are highly non-linear for very small and very large colour differences. We have therefore ignored such scores and taken the intermediate scores to be approximately linear. This has reduced the number of colours considerably, particularly those examined by Thurner and Walther. One point should be noted about visual scores. These go from 1 upwards, but the score for an identical colour should be 0, not 1. If the scoring starts at 1 every score should be reduced by 1 before quantitative comparisons are made with ΔE -values.

Davidson and Friede (1953) produced a quantitative "acceptability" scale by taking the proportion of colour-matchers that accepted or rejected given colours as close matches to standard colours. If visual assessments were consistent and invariable there should of course be a definite limiting value of ΔE in any given direction; such a state of affairs is represented by the full line in Fig. 3. The real results are however probably represented approximately by the pecked line in Fig. 3; in the absence of a series of results in a given direction we have taken an average figure for the slope of the middle portion of the curve. As before we neglected the results at the two extremes.

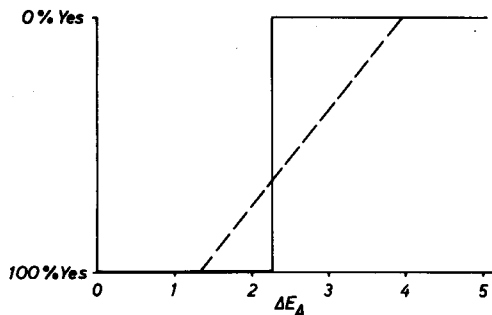


Fig. 3. Hypothetical courses of the acceptability vs colour-difference relation.

Coates *et al* used a ranking procedure in which pairs of colours were compared with a standard, and then the better match from each pair was chosen. The results were quantified by comparing the number of times a colour was chosen with the number of times it would be chosen if the preferences were completely random. This procedure did not yield a value for the colour closest to the standard. We have used the data to calculate the value $V_{A,B}$ but the results appear somewhat uncertain and will not be discussed here.

Schultze and Gall adopted a similar ranking procedure for a series of colours extending in one direction, viz increasing saturation. These were compared with a standard colour of the same hue but differing in lightness. In this way the colour regarded as the best match to the standard was found, and the results were compared with those obtained by the use of colour difference formulae. The value $V_{A,B}$ was not calculated.

2. RESULTS

The following ten colour-difference formulae were compared in every case with the results of visual assessments.

SG	Simon-Goodwin	G mod	Modified Godlove
FMC-2	Friele-MacAdam-Chickering	MR mod	Modified Munsell Renotation
ANS	Adams-Nickerson-Stultz	DIN	
CIE 64		JH	Judd-Hunter
CR	Cube root	SM	Saunderson-Milner

In the modified formulae the value factor was reduced from 4^2 to 2.5^2 .

The data published by Robinson were also treated by several other formulae: the MacAdam-Brown (MB) and Judd-Hunter-Scofield (JHS) formulae, which Robinson himself used, the simple formula investigated by Strocka (1971), and the more complicated formula suggested by Mudd and Woods (1970). The last two formulae are based on differences in x , y , and Y . In the work of Schultze and Gall the original Godlove formula and the Hunter L , a , b ,-formula were also used.

The Table gives $V_{A,B}$ -values calculated from the data of Robinson, Thurner and Walther, Nickerson and Stultz, and Davidson and Friede. Robinson worked with paints, the others all used textiles. Several sets of results obtained by Thurner and Walther, and one set obtained by Davidson and Friede were ignored because the number of specimens was too small. Many sets of results gave high values of $V_{A,B}$ with all ten of the difference formulae; possibly various kinds of error caused a high "noise level", making it impossible to discriminate between the different formulae. We have therefore quite arbitrarily omitted those sets that do not give $V_{A,B}$ less than or equal to 0.30 for at least one difference formula. Thus the Table includes only six sets of results from the 27 sets of Thurner and Walther, three sets of results from Nickerson and Stultz, and twelve sets from Davidson and Friede.

Schultze and Gall investigated a special problem. It had been found that many of the published colour-difference formulae could be expected to give very discrepant results when applied to the evaluation of differences in lightness and saturation of highly saturated reds or blues (Schultze, 1969 *a* and *b*.) When the evaluations obtained by the use of difference formulae were compared with visual assessments of coloured coatings it was found that the agreement ranged from good or tolerable (SG, FMC II, DIN, and JH) through fair (G mod, G, MR mod, CR, ANS, and SM) to bad (Hunter L , a , b) or unspeakable (CIE 64).

For the overall evaluation of colour-difference formulae the results obtained

Table: Comparison of visual scores and colour differences

Series	Standard	Y	No. of col.	SG	FMC 2	ANS	100 . V for formula					MR mod	DIN	JH	SM	MB	JHS	Str	MW
							CIE 64	CR	G mod										
Robinson																			
—	.300	.307	50	25	15.2	21.0	19.0	16.1	16.5	28.5	18.9	37.2	17.2	22.3	14.7	17.2	33.4	41.7	
Thurner-Walther																			
3	.473	.449	39	11	17.9	20.8	16.8	23.3	17.8	18.7	19.7	26.7	23.3	16.2					
7	.381	.325	37	11	45.0	48.4	34.4	46.7	38.2	30.6	26.1	36.2	44.4	40.4					
10	.301	.300	35	17	32.5	36.4	30.6	26.5	30.4	33.1	24.0	41.7	28.4	28.1					
15	.244	.257	7	9	41.2	40.6	36.5	42.9	40.2	42.3	37.1	64.8	26.8	34.6					
24	.268	.271	5	11	38.4	34.1	33.7	35.8	33.4	40.4	31.7	52.2	24.2	30.6					
26	.229	.297	15	11	44.2	38.5	31.8	29.1	31.2	34.6	36.8	35.0	40.9	32.7					
Nickerson-Stultz																			
2	.339	.384	12	25	43.8	37.6	29.3	32.0	30.8	40.9	36.8	38.8	28.5	34.4					
7	.334	.336	6	8	40.7	28.2	39.6	33.4	34.7	53.6	45.1	60.1	16.5	39.7					
9	.349	.364	10	25	57.2	46.2	27.7	38.4	33.8	45.3	38.3	40.6	24.2	34.7					
Davidson-Friede																			
A	.276	.276	18	7	21.5	18.6	19.2	16.8	16.6	20.7	17.9	32.1	12.1	17.5					
D	.302	.271	20	8	26.7	21.2	27.9	19.4	22.5	35.6	21.6	49.0	14.6	24.0					
F	.308	.306	7	8	39.3	40.5	30.0	34.8	33.6	33.5	39.1	74.4	35.6	30.1					
H	.470	.465	41	11	22.1	24.1	39.2	47.4	37.8	36.2	40.5	26.6	25.7	35.3					
I	.344	.333	29	7	17.7	10.3	16.9	12.3	13.5	18.6	18.4	33.9	6.7	18.6					
K	.441	.300	3	6	26.7	15.9	14.7	41.5	19.5	18.0	28.0	35.2	19.7	19.7					
L	.239	.292	3	9	30.1	30.3	21.5	23.1	23.4	23.0	21.4	38.2	28.8	21.2					
M	.540	.333	19	8	40.7	45.4	43.2	73.2	47.2	40.6	25.1	46.6	50.8	47.3					
N	.243	.275	19	9	29.5	27.1	23.5	21.4	21.4	25.5	25.6	24.2	19.0	21.1					
O	.340	.310	38	11	26.0	26.1	18.5	22.1	20.0	23.6	39.2	28.9	21.7	18.0					
P	.473	.471	20	8	19.3	18.8	44.4	31.0	43.6	35.9	24.8	38.3	13.7	35.4					
R	.236	.396	12	13	38.9	33.6	34.1	36.7	29.0	29.2	23.9	47.9	35.0	36.7					

by Robinson should be given particular attention. He carried out only a single experiment, but did this with notable thoroughness. Thirty one colours grouped round a standard—a bluish, somewhat light grey—were assessed four times by 132 people. The closeness of the match was expressed as a score ranging from 1 (“close match to standard”) to 6 (“very poor match”). Figs 4 *a* and *b* give the average scores less 1 as functions of ΔE -values calculated by the SG and DIN formulae.

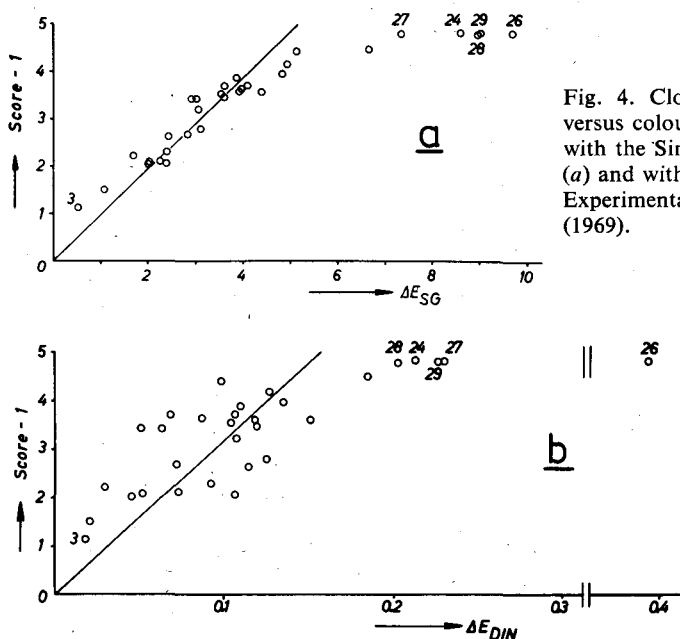


Fig. 4. Closeness-of-match score versus colour-difference calculated with the Simon-Goodwin formula (a) and with the DIN formula (b). Experimental data from Robinson (1969).

The points marked with a number are those that are too close or too far from the standard and have been neglected in calculating $V_{A,B}$. As the tabulated values of $V_{A,B}$ show, the SG and MB formulae give very good agreement with the visual assessments, but the agreement of most of the other formulae is not much worse. However G mod gives poor agreement, and DIN even worse. The formulae investigated by Stroocka and by Mudd and Woods also give poor agreement with visual scores, and it appears that x , y , Y -formulae are in fact of little value.

The DIN formula appears in a poor light in experiments other than that of Robinson: it comes out worst four times in Thurner and Walther's results, once in Nickerson and Stultz's, and six times in Davidson and Friede's. Almost without exception the colours concerned were close to grey or went out into violet, blue, or green. Although the DIN formula gives fair results with the other colours, and turned out well in Schultze and Gall's experiments, it should

be abandoned because of its unreliability with not very saturated colours.

The CIE 64 formula fails with highly saturated colours, particularly blues or reds: this was shown conclusively by Schultze and Gall, and is confirmed by results obtained by Davidson and Friede with their highly saturated red of the series M, the less saturated red K, and the highly saturated yellow H. This formula gives tolerable results in other regions, but it should also be abandoned.

It is difficult to come to any definite conclusion about the other difference formulae. However the JH formula gives good results in general, and in many experiments it turns out better than all other formulae; the only bad results are with Davidson and Friede's series M. The MR mod formula also gives above-average results.

It is interesting to compare the Judd-Hunter and CIE formulae. The projective transformation of the x, y -values is practically the same in each case, and apart from some difficulties with reds this is probably a useful transformation. The basic difference between the two formulae is that the JH formula makes no provision for inclination of the axis; this is clearly better than the excessive inclination of the CIE formula. Another clear advantage with the JH formula is that it gives less weight to differences in lightness compared with differences in chromaticity than the other formulae.

The use of u, v -coördinates is becoming increasingly common, especially in lighting engineering, and it would be interesting to use these coördinates in the JH formula instead of the α, β -coördinates. The combination would have to be carefully tested of course, using available experimental data.

3. SUGGESTIONS FOR FURTHER EXPERIMENTS

Here are a few of my own opinions on how colour-difference formulae might be tested in future experiments.

- (a) The score system for visual assessments of closeness of match seems the most suitable. However a "calibration curve" should be obtained by taking half-a-dozen specimens that depart steadily in one direction from a standard (cf. Fig. 2).
- (b) As many visual assessments as possible, carried out by experienced colour matchers, should be made with 30-40 colours grouped all round a standard. If necessary the same people should repeat their assessments. The measurements should be repeated several times.
- (c) The colours should not quite extend to the limits of the scale of visual scores. These limits can be established by preliminary experiments. Checks should also be made to establish whether or not consistent measurements are made when a colour is very close to the standard.
- (d) For the time being attention should be concentrated on a limited number of regions: a grey, and a highly saturated red, blue, green and yellow.

CONCLUSIONS

When testing the correspondence of visual assessments with colour difference formulae the application of correlation coefficients can lead astray. We use instead a coefficient for mean deviation which has proved to be good when comparing colour difference formulae with each other.

On this base the experimental results of seven papers were evaluated, making use of ten colour difference formulae. On the whole, the DIN formula is so unfavourable in wide colour areas that it should be renounced at. Also the colour difference formula CIE 1964 should no longer be applied as it is definitely unfavourable in the areas of saturated reds up to saturated blues. Among the remaining eight formulae a final choice is scarcely possible for the time being.

Propositions are made for the handling of new test series. The visual assessment by means of a scoring system seems to be the most favourable. A sufficient number of samples (about 30) should be scattered around the standard. Four to six of them should be located in *one* direction from the standard. In this way it would be possible to check the scale of the marks by means of the colorimetric data.

REFERENCES

- Coates, E., and Warburton, F. L. (1968), Colour-difference measurements in relation to visual assessments in the textile field. *J. Soc. Dyers Col.* 84, 467-474
- Coates, E., Day, S., and Rigg, B. (1969), Colour-difference measurements in relation to visual assessments. Some further observations. *J. Soc. Dyers Col.* 85, 312-318
- Davidson, H. R., and Friede, E. (1953), The size of acceptable color differences. *J. Opt. Soc. Amer.* 43, 581-589
- McLaren, K. (1969), Scaling factors in color measurement formulas: A confusing situation. *Color Eng.* 7, 6: 38-44
- Mudd, J. S., and Woods, M. (1970), Colour difference measurement. *J. Oil Colour Chem. Assoc.* 53, 852-875
- Nickerson, D., and Stultz, K. F. (1944), Color tolerance specification. *J. Opt. Soc. Amer.* 34, 550-570
- Robinson, F. D. (1969), Acceptability of colour matches. *J. Oil Colour Chem. Assoc.* 52, 15-45
- *Schultze, W. (1969a), Umfassender Vergleich von sieben Farbabstandsformeln. Proc. 1st AIC Congress "Color 69", Stockholm, 621-640. *Farbe* 18, 105-124
- *Schultze, W. (1969b), Vergleich der Farbabstandsformel nach dem DIN-Farbsystem mit sieben anderen Farbabstandsformeln. *Farbe* 18, 125-130
- *Schultze, W., und Gall, L. (1969), Experimentelle Überprüfung mehrerer Farbabstandsformeln bezüglich der Helligkeits- und Sättigungsdifferenzen bei gesättigten Farben. *Farbe* 18, 131-148
- Strocka, D. (1971), Color difference formulas and visual acceptability. *Appl. Optics* 10, 1308-1313
- Thurner, K., und Walther, V. (1969), Untersuchungen zur Korrelation von Farbabstandsbewertungen auf visuellem Wege und über Farbdifferenzformeln. Proc. 1st AIC congress "Color 69", Stockholm, 671-687. *Farbe*, 18, 191-206

* English translation available

DISCUSSION

McLaren: Correlation coefficients are only of value if they are calculated between ΔE values and some numerical value describing a perceived difference e.g. percent acceptability. The latter is missing from Fig. 1, hence the calculated correlation coefficients are not particularly meaningful. When I used coefficients of variation (a related parameter) I was evaluating scaling factors, not colour difference formulae.

Schultze: Originally, i.e. in my Stockholm paper, the V -formula was introduced to compare two colour difference formulae. Example *a* of Fig. 1 gives, too, a comparison between two colour difference formulae *A* and *B*. The 6 colours are equidistant from S as to the chromaticity, but they are equidistant neither according to formula *A* nor to formula *B*.

As shown later on, there exists, in principle, the same problem when a difference formula is compared with the visual assessment. In case the difference formula *A* or the difference formula *B* is replaced by the visual assessment, it turns out as well that the correlation coefficient does not deliver a correct statement.

Hunter: In the United States, we find that industrial users of color-difference instruments are no longer using measures of total color difference in assessing acceptability of colors of manufactured products. Instead they are relating acceptability to differences in each of the separate trichromatic scales. They will give, for example, separate tolerances in ΔL , Δa , Δb , sometimes even giving separate tolerances in the plus and minus directions for each color scale.

Schultze: This method yields useful details, but it is possible to evaluate from the single values a total difference ΔE and to ascertain the tolerance. In my paper only the total difference is spoken about.

Wyszecki: It is again the problem of acceptability and perceptibility. I am very much interested hearing the various industrial experts on this.

What is exactly the meaning of the figures in the table of your paper? To me all the figures seem to indicate equally bad correlations. I think the only thing you showed is that difference formulae cannot be applied in industrial acceptability work.

Schultze: I would not say that all these figures are bad. One should consider that deviations of the observers' statements and deviations of the measurements in good test series can cause values of about $V = 20\%$. If values below $V = 20\%$ or somewhat above are to be stated, the agreement of the visual assessment with the formula can be called a good one. With some series of Thurner-Walther, Nickerson-Stultz and Davidson-Friede, however, which are not mentioned in this table but which are included in the detailed paper, all the values are above 30% and partly even much higher. That is probably due to a considerable uncertainty of the experiments and I, therefore, have not considered these series when judging the quality of the formulae. It certainly

is not by chance that Robinson's experiments with repeated measurements and a very great number of judgements led to rather low values of V .

Wyszecki: "Bad" is not the right word. We should say, they are not suitable for predicting the industrial acceptability of a mismatch from a standard. Your data do not necessarily show whether they are suitable for predicting the perceptibility of a color difference.

Rigg: Sets where $V_{a,b}$ is more than 0.30 have been omitted. In the practical use of equations the occasions when there is maximum discrepancy between the ΔE values and the visual assessments are the more important ones and therefore it might be better to include all sets.

Schultze: See answer to Dr. Wyszecki.

Clarke: If we look at your table you will notice that there are only three columns with no more than two values over 40: Adams/Nickerson, Saunderson/Milner and Munsell Renotation. I, therefore, wonder what weightings we should give to the rows and the columns.

Schultze: For the reasons I sketched in my answer to Dr. Wyszecki it is not possible to consider the experiments as equivalent. Robinson's experiments are, by far, proved best. Regarding the other tests it is only reasonable to consider each horizontal series on its own and to compare the values V for the various formulae.

Jaekel: Dr. Clarke's question indicates to me that there is a need to search for an overall evaluation, despite the very careful detailed work you have done. Probably there will continue to be a difference in point of view how best to interpret the data: I still feel it is better to look at the overall performance of formulae over many colours first and later that one should look at individual colours more closely.

Schultze: Most interesting to me are the evaluation of the Robinson paper and the paper of Gall and me as it was a special problem. From the other papers (Thurner-Walther, Nickerson-Stultz and Davidson-Friede) with a greater uncertainty of the experiments there can be learnt the following: DIN is bad in many respects, CIE is bad with regard to highly saturated red and blue, for all the other formulae it is not possible to decide which one is better.

Wyszecki: I think there are two steps in getting at a formula describing your industrial acceptance data:

1. The formula under consideration has to be checked with perceptibility data and after that
2. the formula has to be checked with acceptability votings.

Schultze: If we had reasonable perceptibility data, then we could do so. But who provides all these data?

Friele: Dr. Schultze was discussing reliability of available data and mentioned the simple experiment by Robinson to be of outstanding reliability. I remember another single experiment by Balinkin (1939, 1941) which is also of high quality.

REFERENCES

- Balinkin, I. A. (1939), Industrial color tolerances. *Amer. J. Psychol.* 52, 428-448
Balinkin, I. A. (1941), Measurement and designation of small color differences. *Bull. Amer. Ceramic Soc.* 20, 392-402

23 THE UTILITY OF COLOUR-DIFFERENCE FORMULAE FOR MATCH-ACCEPTABILITY DECISIONS*

S. M. JAECKEL

*HATRA, The Hosiery and Allied Trades Research Association
Nottingham, England*

1. INTRODUCTION

Disputes about pass-fail decisions waste time and money. Can quantitative colour-difference (ΔE) limits improve efficiency? CIE-space is visually non-uniform. Colorimeter readings must be transformed into more suitable parameters. Colour-difference formulae emphasize theories of vision and/or perceptibility: they must be tested in commercial situations emphasizing acceptability.

1069 textile patterns in three groups around 12 colour centres provided over 48000 "pass-fail" decisions by experienced industrial assessors and were measured using filter colorimeters. Selection, chiefly to omit almost completely acceptable or unacceptable patterns, left 854 patterns involving over 37000 "pass-fail" decisions. Visual percentage acceptability data and instrumental colour differences for these were correlated statistically. Twenty versions of colour-difference formulae were investigated.

Davidson and Friede (1953) used assessors in one firm and a recording spectrophotometer: investigation of cheaper colorimeters and variation between firms was needed.

This paper is not a full account but concentrates on the overall performance of colour-difference formulae. Further work is in progress.

* The original, of which this is a greatly condensed version, contains additional data and explanatory matter presented at the symposium, and is available from the author.

2. VISUAL INFORMATION

Three groups of knitted false-twist nylon patterns were selected: one dyed at HATRA; one of industrial fabric; and one supplementary group, again dyed at HATRA.

The first group, dyed at HATRA, consisted of 10 sets with 20 plain-knit patterns each; blue gradually going (1) stronger, (2) weaker, (3) redder, (4) yellower; red going (5) weaker, (6) bluer, (7) yellower; and yellow going (8) weaker, (9) bluer, (10) redder. There were three standard patterns, one for each colour, for this HATRA group.

The second group of patterns consisted of industrial fabric for cut-and-sew garments, of 10 sets differing in number of patterns: (000) pink, (300) pale blue, (600) aqua, (700) light fuchsia, (800) deep fuchsia, (900) navy, all in one structure, single piqué (sp), with (100) pink and (400) blue also in "skinny rib" (sr), a structure with pronounced repeating surface irregularity, and (200), (500), in Miralon (M), a lustrous yarn. There were ten standard patterns for this industrial group of patterns, one for each colour in each structure.

After results for these two groups had been obtained, a third group of knitted false-twist nylon patterns was prepared at HATRA, to supplement information from the first group by including different acceptability spacings, higher reflectances, additional colours, and an additional structure, ripple (R), as well as plain-knit. This group consisted of 28 sets with 20 patterns each: blue gradually going (11) weaker, (12) yellower and redder, (13) (R) redder, (14) (R) yellower, and, at half the dye concentration, blue going (15), (16) (R) redder; red going (17) weaker, (18) yellower and bluer, (19) (R) bluer, (20) (R) yellower, and, at half the dye concentration, red going (21), (22) (R) bluer; yellow going (23) weaker, (24) bluer; green going (25) stronger, (26) weaker, (27), (30) (R) bluer, (28), (31) (R) redder, (29), (32) (R) yellower; and brown going (33), (36) (R) bluer, (34), (37) (R) redder, (35), (38) (R) yellower. There were 13 standard patterns, one for each colour at each concentration and in each structure used, for this later group of HATRA-dyed patterns.

The HATRA patterns were mounted on black card protruding from the 95 mm length but not the 80 mm width, the thinner industrial patterns were folded and stapled unmounted to give a four-layer square of side 80 mm. Average viewing distances of 53 cm correspond to 8°–10° subtended at the eye. The assessors' customary viewing conditions varied and included daylight as well as BS950:1967:Part I fluorescent tubes.

Assessors were asked to compare each pattern in turn with the relevant standard pattern and to answer with YES or NO the question "Would you accept this pattern as a commercial match to this standard pattern?" Many provided additional descriptive comments and reasons.

The industrial group and the HATRA-560 group were re-assessed several weeks later, providing 48 assessments per pattern.

For the HATRA-200 patterns, in 1967, there were 32 assessors; 10 from 5 knitting firms (5 from one of the firms) and 22 from 5 dyeing firms (9 from one of the firms). (31 assessors for Set 7, 30 for Sets 8 to 10). For the industrial patterns, in 1969, there were 24 assessors: 8 "standards", from the 32 assessors previously used, for whom they were average in consistency and severity (from five of the ten firms); 10 "suppliers", with various responsibilities, dyers predominating, in the vertical organization providing the patterns; and 6 "buyers", assessors from a retail organization. For the HATRA-560 patterns, in 1970 and 1971, there were 24 assessors: 6 "standard" knitters and dyers, 6 "other" knitters and dyers, 6 "suppliers", and 6 "buyers", altogether representing 12 firms.

The HATRA-200 patterns were selected on the basis of HATRA assessors who proved more severe than the 32 judges from industry, therefore acceptability proved greater than intended. The industrial patterns were selected allowing for the severity level of the HATRA assessor, but the judges were more severe than anticipated and acceptability proved less than intended. Severity differences varied with hue, as shown in Table I

Table I. Acceptability of Colours, First and Second Groups of Patterns.

Number of Patterns: *a* assessed; *b* accepted on less than 50% of occasions; *c* accepted on 5%-95% of occasions

HATRA-200					Industrial				
Sets	Colour	<i>a</i>	<i>b</i>	<i>c</i>	Sets	Colour	<i>a</i>	<i>b</i>	<i>c</i>
(1)-(4)	blue*	79	13	76	000-200	pink	89	50	72
(5)-(7)	red*	59	12	51	300-500	pale blue	69	48	59
(8)-(10)	yellow	60	6	58	600	aqua**	44	38	42
					700&800	fuchsia	65	59	65
					900	navy	42	14	38
TOTAL		198	31	185	TOTAL		309	209	276

* one pattern per colour omitted because of instrument/computer errors

** not included in correlations (See below) for reasons discussed there: omitting this set reduces the total for industrial *c* to 234.

Below the groups are referred to as HATRA-198 (first group), HATRA-185 (those of the HATRA-198 patterns accepted on 5% to 95% of occasions), Industrial (or Ind.)-234 (industrial patterns, aqua set excluded, accepted on 5% to 95% of occasions), and Combined-419 (HATRA-185 and Ind.-234 groups combined).

Of the HATRA-198, 67 (34%) were accepted by less than two-thirds of the assessors and 31 (16%) by less than half. Patterns accepted on less than half the occasions were 24 (13%) for the HATRA-185, 146 (62%) for the Industrial-234, and 170 (41%) for the Combined-419. Fig. 1 provides histograms of visual acceptability for HATRA-185, Industrial-234, and Combined-419 groups.

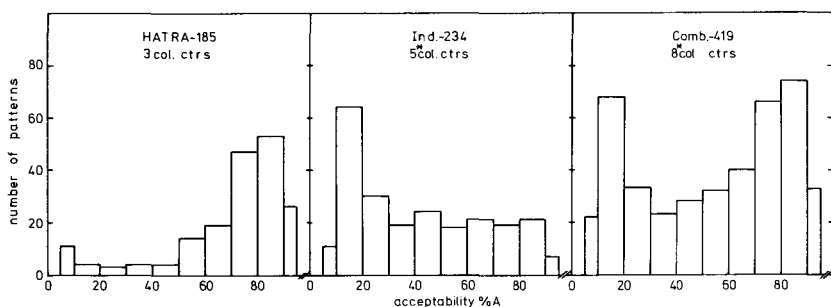


Fig. 1. Frequency Distribution of Pattern Acceptability (*: two fuchsias)

Table II summarizes the hue distribution of acceptability for the third group of patterns and for a balanced selection of plain-knit sets dyed at HATRA from the first and third groups.

Table II. Acceptability of Colours, Third Group of Patterns (HATRA "560") and Selected Plain-knit. Number of Patterns: *a* assessed; *b* accepted on less than 50% of occasions; *c** on 5% - 95% of occasions

Third, HATRA-560, Group						Selected HATRA Plain-knit Group					
Sets	Colour	<i>a</i>	<i>b</i>	<i>c</i>	<i>c</i> Plain	<i>c</i> Ripple	Sets	<i>c</i>	<i>c</i> ₁	<i>c</i> ₂	<i>c</i> ₃ <i>c</i> ₄
(11)-(16)	blue	120	79	91	47	44	(3, 4, 11, 12, 15)	84	25	12	27 20
(17)-(22)	red	120	66	91	46	45	(6, 7, 17, 18, 21)	82	19	11	23 29
(23)-(24)	yellow	40	24	34	34	-	(10, 23, 24)	53	22	5	7 19
(25)-(32)	green	160	89	125	86	39	(25)-(29)	86	25	22	23 16
(33)-(38)	brown	120	67	94	50	44	(33)-(35)	50	16	12	14 8
TOTAL		560	325	435	263	172		355	107	62	94 92

**c*₁ acceptability 5%-25%; *c*₂ >25%-50%; *c*₃ >50%-75%; *c*₄ >75%-95%

Below, the additional groups considered here, all accepted on 5% to 95% of occasions, are referred to as HATRA-435 (those of the third, HATRA-560, group in this acceptability range), HATRA-355 (plain-knit only patterns balanced as far as possible with regard to both hue and within-hue acceptability distribution, selected from the HATRA-185 and HATRA-435 patterns), and Combined-589 (selected HATRA-355 and Industrial-234 patterns combined).

Patterns accepted on less than half the occasions were 236 (54%) for the HATRA-435, 163 (46%) for the HATRA-355, and 309 (52%) for the Combined-589. Fig. 2 provides histograms of visual acceptability for groups 435, 355, and 589.

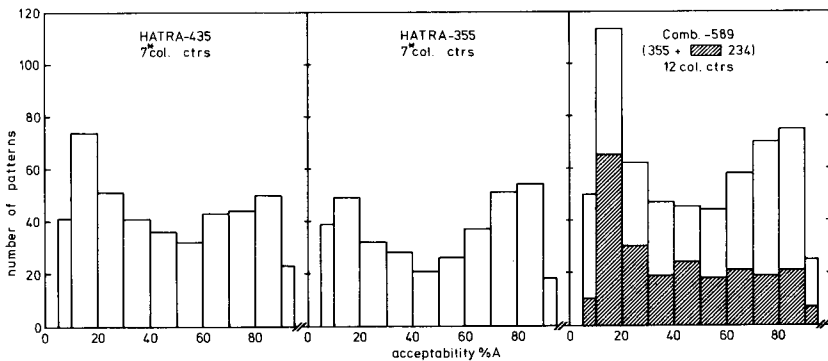


Fig. 2. Frequency Distribution of Pattern Acceptability (*: two depths each of blue and red)

3. INSTRUMENTAL INFORMATION

Preliminary work showed a Colorcord to be less stable than two other makes of filter colorimeter. Prior to visual assessment the HATRA-200 patterns and the associated standards were measured in four orientations in azimuth on a Harrison 1961 RGB_1B_2 Colorimeter and on a Colormaster Mark V. Averaged readings for each filter and pattern were converted by computer first into tristimulus values and then into colour differences according to selected formulae. The industrial patterns were measured similarly on the Harrison RGB_1B_2 only. The HATRA-560 patterns which yielded the 435 group were measured on a different Harrison instrument, the 70 Direct XYZ Colorimeter, in one orientation only. Table III lists CIE specifications of the standard patterns, based on averaged Harrison 1961 readings after every group of five other patterns, for the first two groups, and similarly of Harrison 70 readings for the third group.

Table III. CIE Specifications of Standards

Colour	x	y	%Y	Colour	x	y	%Y
185 HATRA Patterns				234 Industrial Patterns contd. (& aqua)			
Blue	.2024	.2065	19.07	Pale Blue sp	.2732	.2873	59.71
Red	.4270	.2429	18.23	Pale Blue sr	.2677	.2849	59.42
Yellow	.4204	.4515	67.42	Pale Blue M	.2746	.2907	64.79
234 Industrial Patterns				Aqua sp	.2660	.3046	60.92
Pink sp	.3475	.3029	63.89	Light Fuchsia sp	.4588	.2523	22.26
Pink sr	.3494	.3023	61.90	Deep Fuchsia sp	.5218	.2454	12.15
Pink M	.3405	.2992	64.84	Navy sp	.2310	.2027	2.13
435 HATRA: Plain-knit				435 HATRA: Ripple			
Blue	.2128	.1989	19.26	Blue	.2115	.1943	15.89
Half-strength blue	.2274	.2246	27.53	Half-strength blue	.2235	.2155	23.95
Red	.4349	.2448	18.32	Red	.4487	.2446	16.46
Half-strength red	.3998	.2501	25.09	Half-strength red	.4160	.2476	21.64
Yellow	.4084	.4672	70.57				
Green	.2805	.3960	22.84	Green	.2820	.3940	19.91
Brown	.3758	.3146	13.92	Brown	.3807	.3095	11.39

The colour-difference formulae of Table IV were tested.

ACV, Adams Chromatic Value, is expressed with $k=42$, though originally other factors (50 & 43.7) had been used (McLaren (1970a), (1969b)). For Zeta space, $k=1$ to 5 in $\zeta_2=kV_y$ is recommended, though Saunderson & Milner (1946) initially considered 2 to 8. First we used 5(Ze5), later 2(Ze2), when lower lightness weighting proved superior in each of the pairs $M_{\frac{1}{2}} > G_{\frac{1}{2}}I, DH > BB, Fkk > FMC$. Since Ze2 proved better than Ze5, it was optimized, with slight improvements, as found for other formulae by McLaren (1969a) and

Table IV. Colour-difference Formulae and Formula Coding

<i>Code</i>	<i>Reference</i>	<i>Code</i>	<i>Reference</i>
<i>Adams (1942) Spaces</i>		<i>MacAdam (1943) Ellipsoids</i>	
ACV	Nickerson & Stultz (1944)	IR	Ingle & Rudick (1953)
Ze2	Saunderson & Milner (1946)	DF	Davidson & Friede (1953)
Ze5		DH	Davidson & Hanlon (1955)
CUBE ROOT VERSIONS		SG	Simon & Goodwin (1957)
G ₁ ¹ ₁	Glasser et al (1958)	BB	Berger & Brockes (1962)
G ₂ ² ₂	Reilly modified (1963)	FMC	Chickering (1967)
M ₃ ³ ₃	Morton (1969)	Fkk	Wyszecki (1968)
<i>Others, mainly Linear CIE Transformations</i>			
JH	Hunter (1942)	C64	Wyszecki (1963)
SH	Scofield (1943)	Md	Mudd & Woods (1970)
HGf	Hunter (1958)	CDR	Coates, Day, & Rigg (1969)
HGY		ASTM (1968)	

Strocka (1971). Glasser Cube Root was computed both in the original version, G_{11}^1 , and in G_{22}^2 , the modification proposed by Reilly in 1963, which differ as shown in Table V. M_{33}^3 , Morton Cube Root, has $f = 24.8$ and $k = 17$ in $L = f \cdot Y^{1/3} - k$, where G_{11}^1 has 25.29 & 18.38, and CIE1964 Wyszecki has 25 & 17. Morton differentiates to obtain ACV equivalents with “b” term reversed to CIE “hand”, as shown in Table V for M_{33}^3 and ACV with $k = 42$, G_{11}^1 with $K = 41.86$. HGY is also listed in Table V. SH is Scofield’s (1943) “ $7L\alpha$, $7L\beta$ ” modification of JH (NBS unit). Hunter’s (1958) versions HGf and HGY, listed as B_1 and B_2 respectively in ASTM-D2248-68, were criticized by Little (1963), but are used in Hunter and Gardner Color Difference Meters. Y standard was used instead of $\bar{Y}$ mean for e.g. JH(NBS unit) and IR. DF, Davidson and Friede’s (1953) version of MacAdam, is Eq. (1) on p. 584 of their paper. DH, SG, and BB, the three graphical methods applied to the HATRA-185 group, proved time-consuming. Jaekel, Ward, and Hutchings (1963) had used DH for fading of blue light-fastness standards and Berger & Brockes (1962) reported improved performance for this by trebling the DH factor for ΔY , to BB. FMC is the Friele-MacAdam-Chickering formula in Chickering’s (1967), FMC-1, version, Fkk is the FMC-2 modification described by Wyszecki (1968). C64 is the CIE 1964 UVW recommendation. Md, the Mudd Radial Formula, $\Delta E^2 = 1.3^2[200^2(\Delta x^2 + \Delta y^2) + \{\Delta Y/0.0082K(Y + 22)\}^2]$, (for $Y > 8$), uses a K factor linked according to hue to a displaced photopic luminosity curve. It was applied to the Industrial-234 group only. CDR, $\Delta E^2 = 10^5(\Delta x)^2 + [8 \times 10^4(\Delta y)^2 + 0.8(\Delta Y)^2]$, was based empirically on greens.

4. CORRELATION OF VISUAL ACCEPTANCES AND INSTRUMENTAL COLOUR DIFFERENCES

Visual acceptances and instrumental colour differences were correlated in two ways. First, a preliminary correlation was carried out on the HATRA-198 group with “error” scores. Then, the whole body of data was subjected to regression analyses.

a. Error Scores

It was decided to select one set of colorimetric measurements only for de-

ΔE $= [(\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2]^{\frac{1}{2}}$ $\Delta L = L \text{ pattern} - L \text{ standard, etc.};$
 colour —dark —green —blue $L_0 = L \text{ standard, etc.}$
 difference +light +red +yellow $V_x, V_y, V_z: \text{Munsell Value Functions}$
 $Y/Y_{M90} = 1.2219 V_y - 0.23111 V_y^2 + 0.23951 V_y^3 - 0.021009 V_y^4 + 0.0008404 V_y^5$

Formula	ΔL	Δa	Δb
ACV	$9.66 \Delta V_y$	$42 (\Delta V_x - \Delta V_y)$	$16.8 (\Delta V_y - \Delta V_z)$
$M_{\frac{1}{3}}$	$\frac{7.988 \Delta Y}{Y_0^{\frac{1}{3}}}$	$34.717 \left[\frac{\Delta X}{X_0^{\frac{1}{3}}} - \frac{\Delta Y}{Y_0^{\frac{1}{3}}} \right]$	$13.885 \left[\frac{\Delta Y}{Y_0^{\frac{1}{3}}} - \frac{\Delta Z}{Z_0^{\frac{1}{3}}} \right]$
$G_{\frac{1}{3}} 1$	$25.29 \Delta (Y^{\frac{1}{3}})$	$106 \{ \Delta [(1.02X)^{\frac{1}{3}}] - \Delta (Y^{\frac{1}{3}}) \}$	$42.34 \left\{ \Delta (Y^{\frac{1}{3}}) - \Delta \left[\left(\frac{Z}{1.18} \right)^{\frac{1}{3}} \right] \right\}$
$G_{\frac{1}{3}} 2$	$25.29 \Delta (G^{* \frac{1}{3}})$	$106 \{ \Delta (R^{* \frac{1}{3}}) - \Delta (G^{* \frac{1}{3}}) \}$	$42.34 \{ \Delta (G^{* \frac{1}{3}}) - \Delta (B^{* \frac{1}{3}}) \}$
HGY	$10 \Delta (Y^{\frac{1}{3}})$	$17.5 \Delta \left[\frac{1.02X - Y}{Y^{\frac{1}{3}}} \right]$	$7 \Delta \left(\frac{Y - 0.847Z}{Y^{\frac{1}{3}}} \right)$
$G_{\frac{1}{3}} 2: R^* = 1.1084X + 0.0852Y - 0.1454Z$ $G^* = -0.0010X + 1.0005Y + 0.0004Z$ $B^* = -0.0062X + 0.00394Y + 0.8192Z$			

43
 tailed study of colour-difference formulae, to avoid needless repetition. Regression analysis is a less powerful tool, the smaller the number of data, and this selection was carried out separately on the three hues in the first group of patterns dyed at HATRA, since the other two groups discussed in this paper were not yet available. For this preliminary correlation only, therefore, Harrison and Colormaster results were compared on the basis of numbers of rejections. This approach postulates that, if the formula performs similarly to the visual assessor, then it will reject the same number of patterns. If it rejects either fewer or more patterns, its disagreement with visual decisions is bound to be more marked. Hence the assumption, that the formula is to agree with numbers of rejections visually, gives the formula the best possible chance of success, when success is defined as agreement with the current visual situation. Disagreement then is confined to examining to what extent, if the same number of patterns are rejected visually and instrumentally, the rejected patterns are the same patterns in the two cases. To distinguish the ΔE values providing the same numerical pass-fail balance amongst patterns of a particular hue as a particular visual acceptability criterion from those derived by regression analysis, let us call, e.g. for the 50% A level, the former ΔE_{pf50} and the latter ΔE_{50} .

On graphs such as Fig. 3, ΔE against % A was plotted, separately for blues, reds, and yellows, for each formula and colorimeter, with horizontal lines at

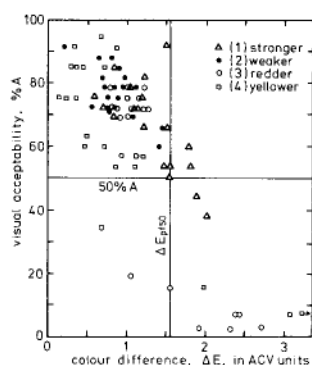


Fig. 3: Visual acceptability ($\%A$) and colour difference (ΔE) according to the ACV formula (see Table V) from measurements on a Harrison 1961 RGB₁B₂ colorimeter for 79 blue patterns in the HATRA-198 group.

the 50% A level. A vertical ΔE_{pf50} line was drawn to the left of as many points as there were below the 50% A line. Points in the upper right and lower left quadrants are errors: $\%A$ and ΔE disagree. Summing for 198 reds, blues, and yellows and dividing by 1.98 gives % wrong decisions for each formula/colorimeter combination (Table VI)

Table VI. Percentage wrong decisions with various formulae and two colorimeters, based on a 50% acceptability criterion

Form- ula	Harr- ison	Color- master
ACV	5	14
SH	7	15
C64	7	15
JH	9	16
IR	10	14
FMC	11	19
Mean	8	16

The real situation—which may be changed by increasing use of instrumental methods—now involves visual decisions. Instrumental decisions based on the Harrison readings agreed better with these in selecting the same patterns for rejection when the same number of patterns were rejected. Therefore Harrison results were twice as realistic as Colormaster ones, so further work was based on the Harrison.

Fig. 3 also illustrates the generally higher acceptability at any particular ΔE level of strength than of hue variation, not unexpected in view of the better performance of formulae with lower lightness weighting, and also observed by Kuehni (1970).

b. Regression Analysis

Nickerson & Stultz (1944), Mudd & Woods (1970), and Strocka (1971) use rank correlation coefficients of the Spearman type, but this does not discriminate between pattern pairs that have similar rank-order differences but widely different quantitative differences (in %A or ΔE), hence makes it difficult to assign instrumental tolerance limits, and may give too optimistic an indication of correlation. This view is shared by Coates, Day, & Rigg (1969). Regression analysis as used here and also by McLaren (1970*b*) postulates a linear relationship between %A and ΔE . For high %A/low ΔE and low %A/high ΔE patterns flattening occurs both in practice (Davidson & Friede (1953)) and in theory (Marshall & Tough (1968)), but in between the approximation is sensible.

For the HATRA-198 analyses, a straight line is assumed for all acceptances. The preponderance of unacceptable patterns in the industrial patterns makes this particularly unrealistic (the sigmoid flattening at the bottom must not be ignored) so analyses are restricted to 234, i.e. 5%-95% acceptance. The aqua set, generally unacceptable, though very similar to the pale blue, clustered about 33%A and 4.4 ΔE (M $\frac{1}{3}$), whereas 33%A for other colours tended to be 2.5 ΔE , so aqua were excluded from all analyses.

The misfit of the aqua set was discovered after all measurements, assessments, and ΔE calculations had been completed. To investigate it further, spectrophotometric reflectance traces were obtained (by courtesy of K. McLaren) for representative selected patterns and their respective standards for the Industrial sets, all supposedly non-metameric like the HATRA-dyed patterns. The patterns examined in each set were non-metameric relative to one another, but relative to their standards aqua were strongly metameric and three sets totalling 75 patterns – (000) pink sp, (100) pink sr, and (400) pale blue sr – were somewhat metameric. Set-specific %A/ ΔE correlations were high or very high for pink sp and pale blue sr. Deletion of these three sets totalling 75 patterns from the Industrial-234 and Combined-419 groups respectively leaves the selected Industrial-159 and Combined-344 groups respectively. %A/ ΔE correlation coefficients (somewhat higher for the best formulae than with the 75 included) provided identical ranking orders of 12 formulae for Industrial-159 and -234 groups on the one hand and for Combined-344 and -419 groups on the other hand. (r decreasing: 159: M $\frac{1}{3}$, G $\frac{1}{3}$ 1, Ze2, ACV, G $\frac{1}{3}$ 2, C64, SH, JH, Fkk, FMC, DF, DH; 344: SH above C64, otherwise ranking same as for 159).

Perhaps because assessors generally used daylight or BS950:1967 Artificial daylight tubes and never tungsten lamps, the presence of a number of somewhat metameric patterns thus has not affected the validity of the results and may indeed increase their relevance to industrial pass-fail situations.

The excluded aqua set may be a misfit not only or mainly because the standard is metameric, but also because it is very much darker (median ΔY 5.8%) than all the patterns and also at one extreme from them in Δx and Δy , the patterns being of low acceptability as well (see Figure 9 below).

In regression analysis, where x (horizontal) and y (vertical) are the two variables and $\bar{x}$ and $\bar{y}$ their mean values, the regression of y upon x (y dependent) minimizes the squares of the vertical displacements (y) from the line and the regression of x upon y (y independent) minimizes the squares of the horizontal displacements. Both kinds of regression were carried out: ΔE upon %A appeared preferable to %A upon ΔE (quoted by McLaren (1970*b*)), because it gives a steeper and hence more realistic slope when %A is plotted vertically (see also the discussion in Section 5 below). In either case, the correlation coefficient r (+1 or -1 if the relationship between the data is exactly a straight line, 0 if there is no relationship

at all) is defined by $[\Sigma(x - \bar{x})(y - \bar{y})]/[\Sigma(x - \bar{x})^2\Sigma(y - \bar{y})^2]^{0.5}$, and is quoted below multiplied by (-100).

Statistical significance of differences in r was tested in three ways. The F test uses the residual variance (σ_e^2) ratio. $\sigma_e^2 \approx [(1 - r^2)\Sigma(y - \bar{y})^2]/(N - 2)$ about the regression line of y (dependent) upon x (N = the number of patterns). It assumes a normal distribution of r , unlikely for very small N values (not the case here) or very high r values (in these results -0.80 is exceeded but rarely). Normal distribution is more assured in Fisher's (1948) z test: $z = \frac{1}{2}[\log_e(1 + r) - \log_e(1 - r)]$. Both tests underestimate significance: they neglect that formulae are intercorrelated, all depending on the same instrument readings. Williams (1959) discusses application of a one-tailed t test using the method of Hotelling (1940): if ri and rii are two $\%A/\Delta E$ correlation coefficients and ro is the corresponding inter-correlation coefficient $\Delta Ei/\Delta Eii$ between the two formulae, $t = (ri - rii)/\{(N - 3)(1 + ro)\}/2(1 - ro^2 - ri^2 - rii^2 + 2roririi)^{0.5}$. Significance increases as ro , and therefore t , increases.

The correlation coefficients for the HATRA-198, HATRA-185, Industrial-234, and Combined-419 groups are listed in Table VII. The flanking lines represent significance groups for three tests, F , Hotelling t , and z . Two of the HATRA-198 patterns had colour differences too large for the graphical methods. The Hotelling t (1940) 5% significance lines are restricted to those best formulae not significantly different according to the z test. Thus they indicate the superior discrimination of the t -test.

Table VII. Correlation coefficients, x (-100), between visual acceptability ($\%A$) and calculated colour difference (ΔE) with various colour-difference formulae. Vertical lines join formulae or values not significantly different at the 5% level according to the F , Hotelling t , and z tests respectively.

HATRA-198(*196)		HATRA-185			Industrial-234			Combined-419(185+234)		
F	z	F	Hot.	z	F	Hot.	z	F	Hot.	z
$M_{\frac{1}{2}}$	79	$Ze2$	77		$M_{\frac{1}{2}}$	75		$M_{\frac{1}{2}}$	82	
$G_{\frac{1}{2}}1$	78	$M_{\frac{1}{2}}$	77		$G_{\frac{1}{2}}1$	74		$G_{\frac{1}{2}}1$	82	
ACV	78	$G_{\frac{1}{2}}1$	77		$Ze2$	73		$Ze2$	81	
$Ze2$	78	ACV	77		ACV	71		ACV	80	
$G_{\frac{1}{2}}2$	75	$G_{\frac{1}{2}}2$	74		HGY	71		$G_{\frac{1}{2}}2$	79	
HGf	74	HGY	73		$G_{\frac{1}{2}}2$	70		HGY	79	
HGY	74	HGf	73		HGf	70		HGf	78	
DH*	73	DH	69		C64	64		SH	73	
$Ze5$	71	JH	68		SH	64		C64	72	
DF	70	DF	67		$Ze5$	57		$Ze5$	70	
IR	70	$Ze5$	66		JH	55		JH	69	
JH	70	IR	66		Fkk	51		Fkk	66	
Fkk	68	CDR	64		CDR	46		CDR	63	
SH	67	SH	64		Md	41		FMC	57	
CDR	67	Fkk	63		FMC	33		IR	55	
SG*	66	BB	62		IR	28		DF	54	
BB*	65	SG	62		DF	26		DH	49	
C64	62	C64	60		DH	17				
FMC	61	FMC	56							

The Combined-419 group gives higher values of the correlation coefficient (r) because of improved acceptance distribution. The Adams spaces (Ze with 2 but not with 5) are always best and HGY is very nearly as good. FMC is always low and the other MacAdam type formulae tried perform badly in the Industrial-234 and Combined-419 groups. Ze5, JH, and SH are intermediate. CDR and Md are rather low. C64 has similar r values in the HATRA-198 and Industrial-234 groups and so is relatively higher in the Ind.-234 list, but is always below ACV. Fkk is not sufficiently better than FMC. Optimizing Ze2 raises r to a maximum of $-.7744$ instead of $-.7726$ for HATRA-185 at Ze 1.3 ($M_{\frac{1}{3}} r = -.7720$) and to one of $-.7263$ instead of $-.7257$ for Ind.-234 at Ze 1.7 ($M_{\frac{1}{3}} r = -.7526$), changes that are somewhat small.

Even the F and z tests confirm the significance of the superiority of $M_{\frac{1}{3}}$, $G_{\frac{1}{3}}1$, ACV, and $G_{\frac{1}{3}}2$ to C64, but Hotelling is much more telling: for the Combined-419 group, the significance of the superiority of $M_{\frac{1}{3}}$ over Ze2 and of ACV over HGY is just at the 5% level, i.e. the chances of the superiorities being fortuitous are 1:20. Other cases are much less likely to be fortuitous. The chances are: for the Combined-419 group, $M_{\frac{1}{3}}$ over ACV and $G_{\frac{1}{3}}2$ 1:2000, over HGY 1:200 000; $G_{\frac{1}{3}}1$ over ACV 1:200, over $G_{\frac{1}{3}}2$ 1:2000, over HGY 1:20 000; for the Ind.-234 group $G_{\frac{1}{3}}2$ and ACV, Ze2, $G_{\frac{1}{3}}1$, and $M_{\frac{1}{3}}$ respectively over C64, 1:200, 1:2000, 1:20 000, and 1:200 000; and for the HATRA-185 and Combined-419 groups, all five best formulae over C64, 1.2×10^9 (for the Combined 419, even on the z test, ACV over C64 was 1:625). Thus for the Combined-419 group, $M_{\frac{1}{3}}$ is significantly superior not only to SH and formulae with lower r , but to all formulae except $G_{\frac{1}{3}}1$.

On these results, Adams Chromatic Value type formulae, especially the Morton Cube Root modification and the original Glasser Cube root, $G_{\frac{1}{3}}1$, should be preferred to all others for accept-reject decisions concerning surface colours such as on textiles. The significance tests confirm that, of the widely used formulae, ACV is preferable to C64; that the simpler $M_{\frac{1}{3}}$ and $G_{\frac{1}{3}}1$ are preferable to ACV; and that the more complex Ze2 offers no advantage over these cube-root formulae.

Part of the procedure applied above was repeated with the HATRA-435 and -355 groups and the Combined-589 group. The HATRA-435 group contains greens and browns, and ripple as well as plain-knit patterns, the HATRA-355 group is a selection of plain-knit only HATRA-dyed patterns from the HATRA-185 and -435 groups that is better balanced both in hue and acceptance distribution, and the Combined-589 group is obtained by considering the Industrial-234 patterns together with the HATRA-355 group. Correlation coefficients are listed with z test significance lines in Table VIII.

The most noticeable feature for the HATRA-435 group is the higher ranking of CDR, JH, and Fkk, relative to their positions for the HATRA-185 and Industrial-234 groups of patterns, and this is largely maintained in the HATRA-355 group. It appears likely that the presence of greens and browns in the HATRA-435 group is responsible for this change.

$M_{\frac{1}{3}}$, $G_{\frac{1}{3}}2$, and $G_{\frac{1}{3}}1$ always appear above ACV. All four of these Adams type formulae always appear above C64 and FMC and are significantly better than these. Formulae very markedly different in efficiency, as indicated by correlation coefficient for different groups of colours assessed under similar conditions, ought to be regarded with great caution and should not be recommended for general use. On this criterion, despite CDR's good showing in the HATRA-355

Table VIII. Correlation coefficients, x (-100), between visual acceptability ($\%A$) and calculated colour difference (ΔE) with 13 colour-difference formulae. Vertical lines join values not significantly different at the 5% level according to the z test (and the Hotelling test for the HATRA-435 group).

HATRA-435		Selected HATRA-355 (92/185 + 263/435)		Combined-589 (HATRA-355 + Ind.-234)	
Hot.	$\bar{z}$		$\bar{z}$		$\bar{z}$
CDR	71	CDR	73	$M_{\frac{1}{3}}$	71
JH	67	JH	70	$G_{\frac{1}{3}}1$	71
Fkk	66	$G_{\frac{1}{3}}2$	70	$G_{\frac{1}{3}}2$	70
$G_{\frac{1}{3}}2^*$	66	Fkk	69	ACV	69
$M_{\frac{1}{3}}$	62	$G_{\frac{1}{3}}1$	68	Ze2	67
$G_{\frac{1}{3}}1$	62	$M_{\frac{1}{3}}$	68	JH	65
ACV	61	ACV	67	HGf	64
Ze2	57	Ze2	64	HGY	63
HGf	52	Ze5	60	Fkk	61
Ze5	52	HGf	59	C64	60
C64	52	HGY	59	CDR	59
HGY	52	C64	56	Ze5	59
FMC	51	FMC	56	FMC	49

Pattern Distribution, %.	Blues	Reds	Pinks	Browns	Yellows	Greens
HATRA-435	20.9	20.9		21.6	7.8	28.7
HATRA-355	23.6	23.1		14.1	14.9	24.2
Combined-589	30.7	25.0	12.2	8.5	9.0	14.6

* According to Hotelling, Fkk does not differ significantly from $M_{\frac{1}{3}}$ and $G_{\frac{1}{3}}1$, whereas $G_{\frac{1}{3}}2$ (between Fkk and the other two) is significantly better than both as a result of the higher values of r_0 , the intercorrelation coefficient. Even on this test for the HATRA-435 group, JH is not better than $G_{\frac{1}{3}}2$; Fkk is not better than $G_{\frac{1}{3}}2$, $M_{\frac{1}{3}}$, and $G_{\frac{1}{3}}1$. Chances of other superiorities being fortuitous include: Ze2 over C64, 1:200; ACV, $G_{\frac{1}{3}}1$, $M_{\frac{1}{3}}$ over C64, 1:20 000; $G_{\frac{1}{3}}2$ and all above over C64, $1:2 \times 10^9$; also $G_{\frac{1}{3}}2$ and $G_{\frac{1}{3}}1$ over ACV, $1:2 \times 10^5$, and $M_{\frac{1}{3}}$ over ACV, 1:200.

group, the very variable CDR and FMC should not be used in practice. The stablest and from this point of view most promising formulae are all of the Adams-Nickerson type, $G_{\frac{1}{3}}2$, the Reilly modification, being outstanding. In the Combined-589 group most of the formulae stable in the constituent HATRA-355 and Industrial-234 groups come to the top in ranking, that is $M_{\frac{1}{3}}$, both Glasser versions of ACV, and ACV itself. (Ze5 of course was equally low in the constituent ranking orders) These additional results do not therefore change the preference for Adams Nickerson Chromatic Value type formulae for accept-reject decisions concerning surface colours, $M_{\frac{1}{3}}$ and $G_{\frac{1}{3}}1$ again appearing most promising, and $G_{\frac{1}{3}}2$ now being ranked above ACV and Ze2 (cf Table VII, Combined-419).

We have applied correlation techniques also to the Davidson and Friede (1953) data. Restriction of these carpet data, all obtained in one firm only (a considerable drawback from the point of view of representative validation, despite the great care exercised in their compilation) to 5%-95% acceptability leaves the

D&F-186 patterns, and correlation coefficients for these involving 11 formulae are presented in Table IX together with the 589 data, separately and combined.

Table IX. Correlation Coefficients, x (-100), between visual acceptability ($\%A$) and calculated colour difference (ΔE) with 11 formulae for the Combined-589, D & F-186, and Overall-775 groups. Vertical lines join values not significantly different at the 5% level according to the z test and the Hotelling-test for the D & F-186 group*

Combined-589 (355 + 234) z		D & F-186 Hot. z		Overall-775 (589 + D & F186) z	
$M_{\frac{1}{3}}$	71	JH	61	$G_{\frac{1}{3}}1$	67
$G_{\frac{1}{3}}1$	71	Fkk	61	$M_{\frac{1}{3}}$	67
$G_{\frac{1}{3}}2$	70	CDR	59	$G_{\frac{1}{3}}2$	67
ACV	69	Ze2	56	ACV	66
Ze2	67	$G_{\frac{1}{3}}2$	54	Ze2	65
JH	65	$G_{\frac{1}{3}}1$	54	JH	64
HGY	63	ACV	53	Fkk	61
Fkk	61	$M_{\frac{1}{3}}$	53	HGY	60
C64	60	C64	53	CDR	59
CDR	59	HGY	48	C64	59
FMC	49	FMC	46	FMC	45

* : $G_{\frac{1}{3}}1$ differs significantly from ACV and HGY, because of high intercorrelation coefficients, r_0 , but not from $M_{\frac{1}{3}}$, C64, and FMC.

The top-ranking formulae for D & F-186, JH, Fkk, and CDR, differ little in correlation coefficient from the values for the Combined-589 group, all the other formulae have lower values for the D & F-186 group. On the z test for the D & F-186 group, the only significant difference is between JH and Fkk on the one hand and FMC on the other. Hotelling results indicate that it is the data that are not very discriminating. Nevertheless, HGY clearly is worse than all Adams type formulae (the chances of this being fortuitous ranging from 1:29 for ACV to $1:2 \times 10^5$ for Ze2) and $G_{\frac{1}{3}}$ is better than ACV (1:22).

The 775 Group consisting of Combined-589 with D & F-186 patterns is the most comprehensive list so far considered. Correlation coefficients and ranking differ hardly at all from the 589 group.

According to z significances, the best formulae, still of the Adams type, do not differ amongst themselves, but $G_{\frac{1}{3}}1$ and $M_{\frac{1}{3}}$ are better than Fkk and all below, and all of these Adams type formulae are better than C64 and FMC. FMC is worse than all other formulae. The conclusions reached already about formula selection stand.

5. RELATIVE PROBABILITIES OF VISUAL AND INSTRUMENTAL WRONG DECISIONS

The probability of wrong decisions was studied, with Blackman's method, for assessors of average consistency but various degrees of severity. The example given is for 50% A for the Industrial-234 group. Method and conclusions differ from those applied to and reached for the Davidson & Friede (1953) data by McLaren (1970b).

Replicate assessments indicate the proportion, p , of wrong assessments for each acceptance decile. To obtain this information, patterns were grouped into visual acceptability heptiles. For each heptile, each assessor's replicate decisions were classified into two acceptances, two rejections, and one acceptance and one rejection per pattern. Each assessor's severity level was found, i.e. the heptile in or after which he switched from more "two rejections" to more "two acceptances" decisions for the patterns in the heptile. The heptiles equidistant from the severity level of each assessor were grouped together for all assessors. For any acceptability group, i.e. distance from the severity level, the proportion, p , of wrong decisions to total decisions was calculated. "Wrong" here means pass where the majority fail, and vice versa. p tends to a maximum of 0.5 at the "switch-over" level for pass to fail, and has a cusp-like distribution (Fig. 4). For the average assessor, of 50% severity, $p = 0.5$ at the average "switch-over" level, i.e. at 50%A. Summing over all deciles the product of $p \times n$ (number of patterns) for each decile gives the total wrong assessments.

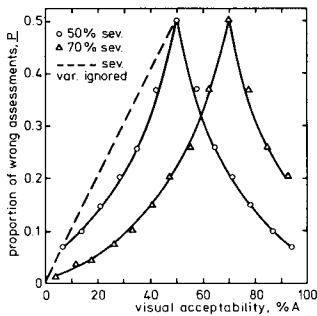


Fig. 4. Error Cusps, 234 Industrial Patterns

The dashed straight line in Fig. 4 indicates the proportion of wrong assessments that would be obtained on the assumption that assessors did not vary consistently in severity. This assumption, although assumed very widely to be correct, need not hold in the general case, indeed it is unlikely, and it was found not to hold in the case of these investigations. The more precise estimation of wrong visual decisions made here by allowing for consistent variation in severity levels amongst assessors, since p values are lowered by the concavity of the cusp, leads to a lower error estimate, relative to the particular "switch-over" level, than is obtained from the straight line assumption implicit in other investigations.

Severity variations were found not only between firms, but also within each of several firms. One and the same firm contained both lenient (14% severity) and severe (86% severity) consistent experienced assessors; i.e., their "switch-over" level from failing to passing was for patterns accepted by more than 14% or more than 86% respectively of all the assessors.

For these assessors, $p=0.5$ at less than (lenient) or more than (severe) 50% A . For deciles intermediate between the population's and these particular assessors' "switch-over" levels, i.e. between 50% A and these assessors' % A with $p=0.5$, what is wrong for the assessor is right for the firm. Hence pattern numbers in these deciles have to be multiplied not by p but by $(1-p)$, the proportion of decisions wrong for the firm although right for the assessor.

The regression line of ΔE upon % A estimates ΔE_{50} . This differs from and is more appropriate here than the $\Delta E_{p,50}$ based on numbers of decisions used in the preliminary correlation with error scores. % A is the independent variable, being the mean of 48 visual judgments per pattern, which can reasonably be taken to represent the true situation of visual decision-making as it is to-day, against which instrumental methods must be judged. Normal curves are centred on the line for ΔE s of each % A decile midpoint. Above 50% A , areas under the curve estimated over ΔE_{50} indicate proportions of wrong instrumental rejections; below 50% A , areas under ΔE_{50} indicate proportions of wrong instrumental acceptances. Summing over all deciles the product of $p \times n$ estimates the number of wrong decisions for that formula. These formula errors are those obtained if all the formulae are set to their equivalents of 50% A . So they differ from the assessor errors, related to severity differences.

The significance of differences between visual and instrumental wrong decisions is shown in Fig. 5. Variances for both "error" populations are derived by $\Sigma np(1-p)$ for all deciles and enable Students' t test to be applied at the 5% level to the corresponding numbers of wrong decisions.

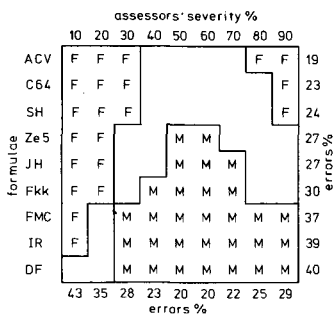


Figure 5. Man or Formula?

Industrial-234 patterns; wrong decisions with instrumental tolerances for different formulae, relative to a single assessor of average consistency and severity. Significantly fewer wrong decisions F by formula, M by man.

ACV, C64, SH, and the right judge are similarly reliable, the other formulae are worse. ACV, to a greater extent than the other three formulae mentioned above, is superior to lenient and severe assessors. It should be noted that C64 performed less well for the 198 group. Since $M\frac{1}{3}$ correlates better with % A than ACV does, it must be superior in this respect also.

The formula errors derived in the above calculations differ from the equation errors discussed in another paper (this symposium, Coates et al 1971) in not being distinct from, but including measurement errors. It is probable that

these latter are of rather lesser importance, but both should be taken into account, as is done here, when visual and instrumental errors are being compared. Since more reproducible measurements can be obtained from newer versions of the instrument used for the Industrial-234 group and since no account has been taken in this comparison of hue-specific variations (see Section 6 below), a cautious and possibly pessimistic view of the efficacy of the instrumental method has prevailed here. Coupled with the favourable view of visual errors obtained by using the cusp, not the straight line, in Fig. 7, this makes the conclusion even more compelling that, provided the right formula is selected, the instrumental method is at least as good in some circumstances and better in others than the existing visual situation.

6. PROBLEM AREAS*

The Harrison 1961, though more reliable in this instance than a Colormaster Mark V, may not be reproducible enough for production control, but a HATRA-modified Harrison 70 Direct XYZ Colorimeter is.

Not only assessors, but also assessor groups, vary in severity with colour and direction of change from standard. Tolerance-limit variation with severity and colour is more marked for the worse formulae, especially for browns and navys relative to other hues. All formulae can give mediocre results for some groups of patterns if the components of ΔE are ignored. Choice of independent

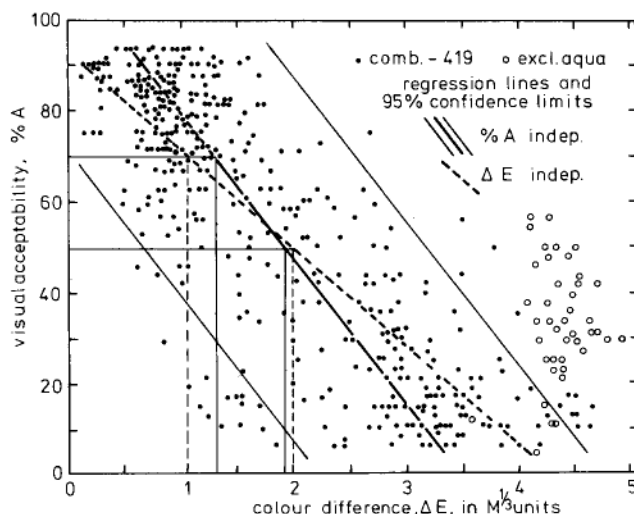


Figure 6. Visual Acceptability (%A) and Colour Difference (ΔE) according to the Morton Cube Root Formula and the Influence of Choice of Independent Variable on Corresponding Values from the Regression Lines.

* This section in particular has been greatly abridged.

variable, indicated in Fig. 6, can have a more marked effect on tolerance limits than assessor-group variation.

7. CONCLUSIONS

When over 37000 visual "pass-fail" decisions for three groups of textile patterns totalling 854 and involving 12 colour centres were correlated by regression analysis with colour differences computed for up to 20 colour-difference formulae, those formulae based on Adams Chromatic Value performed significantly better both than most other formulae and than assessors too severe or too lenient, and performed at least as well as assessors of correct severity. Therefore these formulae are to be preferred to others for pass-fail decisions involving surface colours such as dyed textiles. The Morton and Glasser Cube Root formulae performed best. When the Davidson & Friede data were also included, they remained superior. Problems requiring attention to ensure successful application in industrial control situations include taking cognizance of instrument reliability as much as of reproducibility, of assessor and hue variation of tolerance limits, of choice of regression line, and of components of ΔE , the colour difference.

8. ACKNOWLEDGMENTS

To the Director and Council of HATRA for permission to publish and to the Ministry of Technology for financial support, and, with gratitude, to the co-workers at different times with various stages of these investigations, F. Blackman, C. D. Ward, B. J. Hojiwala, Mrs. Jill Lockett, and Mrs. Zella Griffiths, assisted by D. Harding, Miss Barbara Holehouse, Miss Norma Shepherd, and Miss Christine Walden, without whose efforts there would have been no paper.

9. REFERENCES

- Adams, E. Q. (1942) X-Z planes in the 1931 ICI system of colorimetry. *J. Opt. Soc. Amer.* 32, 168-173
- American Society for Testing and Materials (1969) D2244-68 Standard method for instrumental evaluation of color differences of opaque materials. Pt 21—Paint, 463-471 Philadelphia
- Atherton, E., Garrett, D. A., & Vickerstaff, T. (1954) Realisable colour gamuts in dyeing. *J. Textile Inst. (Proc.)* 350-368
- Berger, A. & Brockes, A. (1962) Zur Beurteilung des Ausbleichens von Lichtechtheitstypen mit Farbdifferenzformeln. *Farbe* 11, 263-274
- Chickering, K. D. (1967) Optimization of the MacAdam-modified 1965 Friele color-difference formula. *J. Opt. Soc. Amer.* 57, 537-541
- Coates, E., Day, S., & Rigg, B. (1969) Colour-difference measurements in relation to visual assessments—some further observations. *J. Soc. Dyers Col.* 85, 312-318

- Coates, E., Day, S., Provost, J. R., & Rigg, B. (1971) Colour-difference equations for setting industrial colour tolerances. This Symposium
- Davidson, H. R. & Friede, E. (1953) The size of acceptable color differences. *J. Opt. Soc. Amer.* 43, 581-589
- Davidson, H. R. & Hanlon, J. J. (1955) Use of charts for rapid calculation of color difference. *J. Opt. Soc. Amer.* 45, 617-620
- Dinsdale, A. & Malkin, F. (1971) Colour tolerances of ceramic wall tiles. This Symposium
- Fisher, R. A. (1948) Statistical methods for research workers. Oliver & Boyd. London
- Glasser, L. G., McKinney, A. H., Reilly, C. D., & Schnelle, P. D. (1958). Cube-root color coordinate system. *J. Opt. Soc. Amer.* 48, 736-740
- Hotelling, H. (1940). The selection of variates for use in prediction with some comments on the general problem of nuisance parameters. *Annals of Math. Stat.* 11, 217-283
- Hunter, R. S. (1942) Photoelectric tristimulus colorimetry with three filters. *J. Opt. Soc. Amer.* 32, 509-538
- Hunter, R. S. (1958) Photoelectric color difference meter. *J. Opt. Soc. Amer.* 48, 985-995
- Jaekel, S. M., Ward, C. D., & Hutchings, D. M. (1963) Variations in assessment of light fastness and in rates of fading and spacing of the blue standards. *J. Soc. Dyers Col.* 79, 702-714
- Kuehni, R. (1970) The relationship between acceptability and calculated color-differences on textiles. *Color Engng.* 8: 1, 47-53
- Little, A. C. (1963) Evaluation of single-number expressions of color difference. *J. Opt. Soc. Amer.* 53, 293-296
- MacAdam, D. L. (1943) Specification of small chromaticity differences. *J. Opt. Soc. Amer.* 33, 18-26
- McLaren, K. (1969a) The precision of textile colour matchers in relation to colour difference measurements. Proc. 1st AIC congress "Color 69" Stockholm, 688-708. *Farbe* 18, 1/1-190
- McLaren, K. (1969b) Scaling factors in color difference formulas: a confusing situation. *Color Engng.* 7, 6: 38-45
- McLaren, K. (1970a) The Adams-Nickerson colour-difference formula. *J. Soc. Dyers Col.* 86, 354-366
- McLaren, K. (1970b) Colour passing—visual or instrumental? *J. Soc. Dyers Col.* 86, 389-393
- Marshall, W. J. & Tough, D. (1968) Colour measurement and colour tolerance in relation to automation and instrumentation in textile dyeing. *J. Soc. Dyers Col.* 84, 108-119
- Morton, T. H. (1969) ΔE by differential colorimetry in production dyeing. Proc. 1st AIC congress "Color 69" Stockholm, 709-716. *Farbe* 18, 164-170
- Mudd, J. S. & Woods, M. (1970) Colour difference measurement. *J. Oil Colour Chem. Assoc.* 53, 852-875
- Nickerson, D. & Stultz, K. F. (1944) Color tolerance specification. *J. Opt. Soc. Amer.* 34, 550-570
- Saunderson, J. L. & Milner, B. I. (1946) Modified chromatic value color space. *J. Opt. Soc. Amer.* 36, 36-42
- Schultze, W. (1971) The usefulness of colour difference formulae for fixing colour tolerances. This Symposium
- Scofield, F. (1943) A method for determination of color differences. *Natl. Paint Varnish Lacquer Assoc., Sci. Circ.* 664
- Simon, F. T. & Goodwin, W. J. (1957) Rapid graphical computation of small color differences. *J. Opt. Soc. Amer.* 47, 1050
- Strocka, D. (1971) Color difference formulas and visual acceptability. *Appl. Optics*, 10, 1308-1313
- Williams, E. J. (1959) Regression analysis. John Wiley & Sons, New York
- Wyszecki, G. (1963) Proposal for a new color-difference formula. *J. Opt. Soc. Amer.* 53, 1318-1319
- Wyszecki, G. (1968) Recent agreements reached by the colorimetry committee of the CIE. *J. Opt. Soc. Amer.* 58, 290-292

DISCUSSION

Coates: It should be pointed out that the CDR equation, included in your tests for goodness of fit to your data, was obtained solely on data for series of green samples. It was never intended nor expected to provide a good fit in other regions of colour space. It is gratifying to find that it does provide a good fit to some of your data.

Billmeyer: The correct identification for your Fkk formula is FMC-2.

Jaekel: Fkk as defined in my paper and FMC-2 as characterized elsewhere refer to the same formula, but neither identification is more "correct" than the other: FMC-2 has been customary, Fkk as defined by me is identified correctly. I feel the proliferation of lengthy abbreviations is to be deplored and that three typing spaces can suffice to identify all colour-difference formulae and to distinguish them unambiguously from each other as I hope to have shown in the codes I use. With "FMC-1" and "FMC-2" one has to think: "which one has the two "k" factors?", with Fkk this is immediately obvious, and in three, not in five, spaces. Fkk, I should like to suggest, is not only just as correct as FMC-2, but also better!

Bartleson: Early in your preprint you mentioned that there were some problems associated with the instrumentation. Can you summarize what it was that you were referring to?

Jaekel: The instrument we used was good enough for comparing colour-difference formulae when all patterns were measured by one observer at one session with the relevant standards being remeasured at appropriate short intervals. That 1961 version of the Harrison was not robust and good enough for a production control situation. The version of the Harrison 1970—not available when our work started—that we have decided to use now and that we incidentally used for the HATRA-435 patterns is much better in that respect and will do for the control situation.

We in fact compared the Colormaster and the 1961 Harrison, and for the HATRA-185 data the Colormaster did not agree as well with the visual judgments, so we used the Harrison data for the further comparisons. A fuller description than could be provided in the preprint will be in the version of the paper published.

In my opinion the probability of making more or fewer wrong decisions than visually, depending which formula is used, is more important though than the particular instrument used, provided the instrument used was adequate.

Bartleson: Your point is well taken that there are several sources of variability, the specification of the physical characteristics being one of them.

Jaekel: I have some slides on that, bearing on the variability of instruments and the variability of people in making "pass-fail" decisions. They show how we worked out the relative probabilities of wrong decisions in the paper—to

be given in slightly greater detail in the published version than in the preprint —, and stress the importance we attach to differences in severity of assessors. We may be publishing a full account of this approach elsewhere.

Strocka: 1. Do you plan to construct from your data tolerance ellipses?
2. Do you plan to make your data available?

Jaeckel: 1. We can consider this, but we concentrated in the HATRA dyeings on colour changes in a few specific directions from each standard pattern, so ellipse spaces might not be filled sufficiently randomly, representatively, and reliably. A few of the Industrial-234 group might possibly be suitable, but not all, because some directions of difference from standard happen to be predominant in these more randomly selected patterns. However, we had a very quick look at what kind of ΔE_{50} levels we get for the different formulae on a colour-specific basis. The way we did it may not be acceptable to everybody since it is the straight line regression with %A independent — which we think we ought to prefer, for several reasons.

2. Yes, I definitely have every intention to do so eventually, possibly in a series of papers elsewhere.

24 COLOUR-DIFFERENCE EQUATIONS FOR SETTING INDUSTRIAL COLOUR TOLERANCES*

E. COATES, S. DAY, J. R. PROVOST AND B. RIGG

University of Bradford, Bradford, Yorkshire, England

INTRODUCTION

A colour difference equation which successfully predicted the colour differences between similar colours would be of immense value in industrial tolerance/shade passing work. Various investigations have been carried out by Nickerson and Stultz (1944), Davidson and Friede (1953), Thurner and Walther (1969), Coates and Warburton (1968) in attempts to assess the suitability of existing equations for use in such work. In general the correlation obtained between the calculated ΔE values, and the mean visual assessments was poor. It is well known (Coates, Provost and Rigg, 1970) that visual estimates of colour differences vary according to the conditions used for assessment. Thus, different uniform lightness scales are obtained when series of grey samples are assessed against different backgrounds (Judd and Wyszecki, 1963), while Brown (1952, 1957) has shown that chromaticity discrimination ellipses depend on the colour of the background and on the observer. The conditions used in testing a colour difference equation should, therefore, correspond fairly closely to those appropriate to the practical application. For the shade passing of surface colours the samples should be large, corresponding to a field size of 10° or more, and the colour differences small, say less than 3 CIE units.

Most colour difference equations appear to be based on information such as that contained in the Munsell Renotation System (Newhall, Nickerson and Judd, 1943) or on information obtained using visual colorimeters such as the MacAdam (1942) chromaticity discrimination ellipses. The Munsell system is based on visual judgements of relatively large colour differences with individual samples subtending an angle of 2° and with a neutral background around and between the samples. The MacAdam ellipses are based on a 2°

* Read by B. Rigg

field of view and a single observer. These sets of data do not provide all the information required to derive a colour difference equation. For any particular point in Munsell colour space the important factors are the relative sizes of the units of hue, value and chroma, information which is not included in the system. Attempts to determine these factors have provided widely differing results (Godlove, 1951). Data obtained using visual colorimeters provide no information regarding the possible variation of chromaticity discrimination with luminance factor.

TESTING COLOUR — DIFFERENCE EQUATIONS

(1) Davidson and Friede data.

In testing colour difference equations one may wish to know (a) which of several equations gives ΔE values in closest agreement with the visual assessments and (b) how accurate the ΔE values from a particular equation are (possibly compared with the accuracy of visual assessments by a single observer). The largest single body of experimental data appropriate to industrial conditions is that obtained by Davidson and Friede (1953). This data has been reanalysed to try to answer the above questions, particularly (b). Approximately 40 individual visual judgements were made on each of 287 colour differences by eight observers who were asked to accept or reject each sample in turn as a commercially acceptable match to the appropriate standard (19 in all). The results were presented in the form of the percentage acceptance (%A) of each sample. The relative merits of several colour difference equations were assessed by inspection of plots of %A against ΔE .

This approach is rather subjective and not suited to assessing the accuracy of the ΔE values. It is obviously desirable to scale the %A values to values (ΔV) which are directly proportional to the ΔE values. This was done by the procedure described previously (Coates, Day and Rigg, 1969). Samples for which %A was not within the range 10 to 90 were excluded from the analysis. A plot of ΔV against ΔE should be linear with the scatter of the experimental points decreasing for equations giving more accurate ΔE values. However, even for a perfect equation some scatter would remain because of errors both in the experimental measurement of the tristimulus values of the samples and in the visual assessments. Allowance was made for these errors in calculating the equation error from the experimental scatter. The measurement error (expressed as a standard deviation) was taken to be 0.1 CIE units. This is roughly consistent with Davidson and Friede's claim that the standard deviation of the measurement error was less than 0.0005 in x or y . The error in the visual assessments arises because of the use of a finite number of observers. For a series of n trials where the probability of success is p , the number of successes is distributed according to a binomial distribution with a mean of np and a

standard deviation (σ) of $\sqrt{np(1-p)}$. Thus, for a sample judged 40 times where $\%A = 60$, $\sigma = \sqrt{40 \times 0.6 \times 0.4} \sim 3.1$.

This leads to a value of ~ 0.2 for the standard deviation of the ΔV values.

Various possible measures of the equation error were considered. Since the units of both ΔE and ΔV are arbitrary some factor (f) must be introduced such that ΔE is on average the same size as $f\Delta V$. From the form of colour-difference equations it would appear that errors will in general be proportional to the ΔE values. If a particular constant in an equation is 10% high this will cause the same percentage error in a small colour difference as in a large one. However, in the case of a poor equation, for a true difference of say 2 the equation could give a ΔE of 4 or more but not zero or less. Thus, large positive errors are more likely than equally large negative ones. The best simple solution appears to be to express the equation error as $\sigma(\log \Delta E)$ where $\sigma(\log \Delta E) = [\Sigma(\log \Delta E - \log f\Delta V_i)^2/n]^{1/2}$ where n is the number of colour differences and ΔV_i is the true visual difference, i.e. based on the assessment of an infinite number of observers. Hence, if $\sigma(\log \Delta E) = 0.18$, then for a one standard deviation error the two limits for a ΔE of 3.6 are 5.4 and 2.4, i.e. 3.6 multiplied or divided by 1.5.

For any particular colour difference equation $\sigma(\log \Delta E)$ was calculated from the scatter on the ΔV against ΔE plot; the total error for any experimental point was considered to be due to a combination of the three separate sources of error. The slope and intercept of the best straight line were taken to be those consistent with the minimum value of $\sigma(\log \Delta E)$. A detailed account of the calculation has been given elsewhere (Coates, Day, Provost and Rigg, 1972). The results obtained for various colour difference equations are given in Table I. The equation errors are expressed as the standard deviation of $\log \Delta E$ and also as the coefficient of variation (CV) of ΔE .

A CV of 40 implies that the 1σ limits on a ΔE of 4 are $4 \pm 0.4 \times 4$ i.e. 2.4 and 5.6. The CV values were estimated from the $\sigma(\log \Delta E)$ values and must be regarded as approximate values only.

Table I. Equation Errors from Davidson and Friele Data.

Equation	CIE	GCR	MFC2	AN(42)	JH	Eqn 1
$\sigma(\log \Delta E)$	0.22	0.20	0.18	0.20	0.16	0.14
CV	46	42	40	42	36	32

The 1964 CIE, Glasser Cube Root and MacAdam – Friele Chickering equations were used in the form recommended by the CIE (1967). The Adams – Nickerson and Judd – Hunter equations were used in the form given by Coates and Warburton (1968). Eq. 1 is given below.

The order of merit of the equations is the same as that obtained from the correlation coefficient (Coates, Day, Provost and Rigg, 1972). Small changes

in the values chosen for the errors in the instrumental measurements and visual assessments had very little effect on the calculation. The equation errors seemed very large and therefore the possibility of finding an equation giving a better fit to the data was investigated. One of the best (and the simplest) equations found was

$$\Delta E^2 = (\Delta L)^2 + 0.5 (\Delta C)^2 \quad (1)$$

where $(\Delta C)^2 = [10^5 (\Delta x)^2 + 10^5 (\Delta y)^2] Y_s^{\frac{1}{2}}$.

$L = 50 \log Y$ and Y_s is the luminance factor for the standard colour (on a scale from 0 to 100%). This equation is denoted by Eq. 1. in Table I. According to the significance test described by Guilford (1965) this equation fits the data better (0.1% level) than any of the other equations. This result is surprising particularly since Eq. 1 involves no transformation of the x, y chromaticity diagram.

Obviously any conclusions are valid only in so far as the experimental data are representative of colour-differences in general. The Davidson and Friede samples appear to be fairly well distributed round the standard colours used by them. The latter are plotted on a chromaticity diagram in Fig. 1. The chromaticity gamut covered is only a small fraction of the total gamut available,

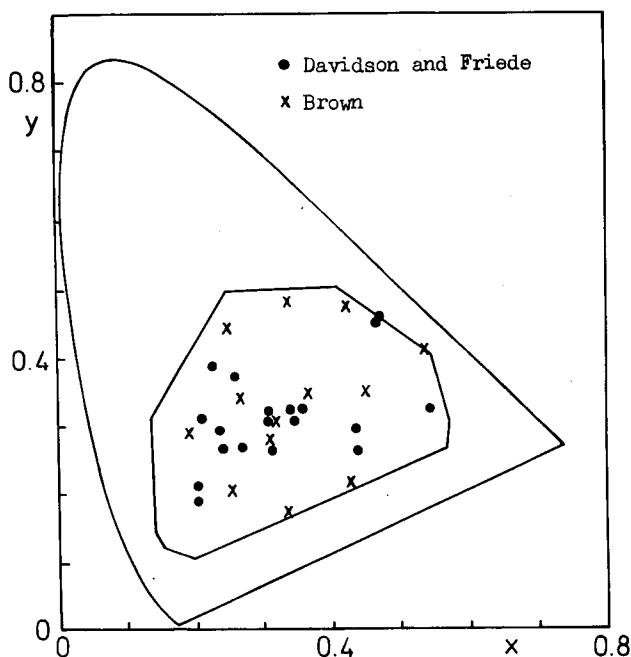


Fig. 1. Colour centres in relation to the chromaticity gamut realisable with surface colours.

but must be compared with that actually obtainable with surface colours. The solid line gives the boundary determined by Atherton, Garrett and Vickersstaff (1954) for real dyestuffs on a variety of substrates. Although the boundary for pigments is slightly different and the gamut for dyestuffs will be somewhat greater at the present time, the coverage provided by the Davidson and Friede samples appears to be reasonably satisfactory.

(2) Brown Ellipses.

The chromaticity discrimination implicit in Equation 1 and other colour-difference equations was tested further using the chromaticity cross-sections of the colour-matching ellipsoids determined by Brown (1957). Brown used a visual colorimeter in which the colour under investigation occupied a 10° field of view. For all points on the perimeter of the ellipses the colour differences from the centres should be constant. Using any one colour difference equation ΔE values were calculated for 8 points at 22.5° intervals around the perimeter of each ellipse by a procedure very similar to that used by MacAdam (1964). Of the 22 ellipses only the 13 which fell within the gamut realisable with surface colours were used. The positions of these are indicated in Figure 1. The ellipses corresponding to the weighted averages of the 12 observers were used in the calculations. The $\sigma(\log \Delta E)$ values were calculated directly, no account being taken of errors in the measurements or visual assessments. The results are given in Table. II.

Table II. Equation Errors from Brown Ellipses.

Equation	CIE	GCR	MFC2	ACV	JH	Eqn 1
$\sigma(\log \Delta E)$	0.18	0.16	0.18	0.17	0.15	0.13
CV	40	34	39	37	34	30

The pattern is very similar to that obtained from the Davidson and Friede data. The equation errors appear to be slightly less, presumably due to the colour differences being confined to chromaticity differences in our analysis of the Brown data. None of the equations gives a very good fit, but Eq. 1, again gives the best fit. It must be stressed that the two sets of experimental data were obtained by completely different methods. The Davidson and Friede observers were asked to consider the acceptability of the colour differences while the Brown observers were concerned with the perceptibility of the differences. It seems reasonable to conclude that for small colour differences between large areas of surface colours, Eq. 1 agrees better with visual assessments than the other equations studied and would therefore be more suitable for industrial use. It is intended to test Eq. 1 with further experimental data as this becomes available. In view of the simple form of the equation and the poor fit to visual assessments (in absolute terms) it seems highly likely that

equations giving better fits should be obtainable and work designed to achieve this end is in progress.

The *Zeta-Eta* chromaticity space recently recommended by MacAdam (1971) has also been studied. For the 13 Brown ellipses values of σ ($\log \Delta E$) = 0.13 and $CV = 31\%$ were obtained. With the Davidson and Friede data the *Zeta-Eta* chromaticity space was combined with luminance factor differences in several different ways the best results were obtained with a form similar to Eq. 1. when values of σ ($\log \Delta E$) = 0.15 and $CV = 34\%$ were obtained. These results are almost as good as those for Eq. 1. but are disappointing in view of the complexity of the *Zeta-Eta* space.

ACCURACY OF VISUAL ASSESSMENTS.

The performance of colour difference equations must be compared with the accuracy of the visual assessments made by a single observer since colour tolerances are normally judged by a single colourist. The only published estimate of visual accuracy in judging colour differences appears to be that obtained by McLaren (1970). The work was confined to colours varying in only two dimensions. The accuracy varied with the size of colour difference but for differences of the order of a commercial tolerance limit (~ 1.5 CIE unit) the errors corresponded to $CV \sim 30\%$. It is also possible to obtain estimates from studies not directly concerned with visual accuracy. Thus in our analysis of the Davidson and Friede results, the slope of the ΔV against ΔE plots is a measure of the visual accuracy. The unit of ΔV is the standard deviation of a visual assessment. Thus if the slope is f , 1 visual standard deviation corresponds to $1/f \Delta E$ units. If the mean ΔE is $\overline{\Delta E}$, then the accuracy of a single visual assessment expressed as a coefficient of variation is given by $100/f \overline{\Delta E}$. The value obtained varied slightly depending on the particular colour difference equation used but averaged about 50%. This value is probably more realistic since it is based on assessments where the colour difference was compared with an imaginary tolerance limit, the particular judgement which has to be made in practice. Trial calculations have shown that if different observers worked to different tolerance limits this would have little influence on the $\Delta V/\Delta E$ plot. Further estimates have been obtained from other work including some where the colour differences were compared by the method of ratio comparisons. (Coates, Provost and Rigg, unpublished work) where the accuracy of a single observer could be obtained directly. Although the values varied considerably, $CV = 50\%$ was found to be a reasonable average for the size of colour difference under consideration (ΔE 1–2 CIE unit). For smaller differences CV increases.

COMPARISON OF EQUATIONS WITH VISUAL ASSESSMENTS

Comparison of the accuracy of visual assessments with the values for equation accuracy given in Tables I and II suggests that most colour difference equations give ΔE values slightly more reliable than the visual assessment of one colourist. However, if there was much error in the instrumental measurements the overall instrumental plus equation error could easily be greater than that visual error.

Calculations have been carried out in an attempt to assess the consequences of using instrumental methods of shade passing instead of visual methods. The basic approach was similar to that outlined by Marshall and Tough (1968). A manufacturer of a coloured article works to a certain tolerance limit. However, because of errors in the visual assessments of his colourist, some samples above the tolerance limit are considered acceptable and sent to the customer, possibly a retail organisation. The samples are re-assessed by the retailer who obviously works to a considerably higher tolerance limit since, despite his errors, very few samples are returned to the manufacturer. If the manufacturer were to use instrumental procedures based on a colour difference equation ($CV = 40\%$) rather than visual assessments ($CV = 50\%$) then with the same tolerance limit fewer complaints would be expected from the retailer. Alternatively, if the level of complaints was satisfactory, the tolerance limit could be raised slightly, resulting in fewer samples being reprocessed. In either case there would be direct or indirect economic benefit to the manufacturer which would be greater the better the colour difference equation. The calculations indicate that, allowing for reasonable instrumental errors, the advantages of using existing equations would be relatively slight but much more benefit could be obtained using an equation with an error (CV) of 25% or less.

If the manufacturer and retailer were to agree on the use of a particular colour difference equation then the only source of disagreements between the two would be due to errors in instrumental measurements and these should be comparatively low. It would thus be possible for the manufacturer to increase his tolerance limit, the retailer to decrease his limit and at the same time reduce disagreements between the two. However, this procedure ignores the fact that the retailer sells the product to the general public who will judge the colour visually. Calculations show that if the manufacturer uses a particular equation and tolerance limit then less complaints can be expected from the customer (general public) if the retailer assesses the samples visually rather than by using the same equation. This result arises because the retailer by using visual estimates will reject some of these samples which his customers would reject, while the instrumental method merely checks the manufacturers measurements. Thus, the use of the same equation by manufacturer and retailer could be a definite disadvantage to the latter. It is of course possible that in practice manufacturer/

retailer disagreements are more serious than those between the retailer and his customers or that benefits accruing to the manufacturer could be passed on to the retailer in the form of reduced prices. In either of these cases the use of the equation could be beneficial to the manufacturer and the retailer.

CONCLUSIONS

It is of extreme importance that the conditions e.g. field size, background colour, the use of average observer data and the size of the colour gamut for which an equation is established should correspond to the conditions under which the equation is to be applied. If an equation is to be accepted for general use in industry, that equation should be the best one available for the conditions applying in industry (large samples and small colour differences). The fact that most equations give a poor fit (in absolute terms) with visual assessments made under these conditions is probably due to the equations being designed to fit other conditions.

The best of the equations used in this paper when applied to data obtained under industrial conditions (Eq. 1) is a simple empirical equation essentially based on the x, y (CIE) chromaticity diagram. The same equation also gives the best fit to discrimination data obtained using a visual colorimeter and a 10° field of view when the data are restricted to the colour gamut obtainable with surface colours. This implies that for these conditions the chromaticity transformations involved in the other equations produce planes less visually uniform than the original x, y plane.

Estimates of the errors likely to be made in assessing colour differences by a single observer suggest that these are even larger than the equation errors. For shade passing in industry it seems likely that some benefit could be obtained by using almost any one of the suggested colour difference equations. The least benefit would probably be obtained by the use of the 1964 CIE equation while the benefits would increase sharply the better the equation, at least until the equation error became comparable with the measurement error. However, it could be unwise for a retailer to agree with a manufacturer on the use of any existing colour difference equation.

REFERENCES

- Atherton, E., Garrett, D. A., and Vickerstaff, T. (1954) Realisable colour gamuts in dyeing. *J. Textile Inst.*, 45, 350.
- Brown, W. R. J. (1952) The effect of field size and chromatic surroundings on color discrimination. *J. Opt. Soc. Amer.*, 42, 837-844.
- Brown, W. R. J. (1957) Colour discrimination of twelve observers. *J. Opt. Soc. Amer.*, 47, 137-143.
- Coates, E., and Warburton, F. L. (1968) Colour—difference measurements in relation to visual assessments in the textile field. *J. Soc. Dyers. Col.*, 84, 467-474.

- Coates, E., Day, S., and Rigg, B. (1969) Colour-difference measurements in relation to visual assessments—some further observations. *J. Soc. Dyers Col.*, 85, 312-318
- Coates, E., Provost, J. R., and Rigg, B. (1970), Colour-difference measurements in relation to visual assessments—effect of viewing conditions. *J. Soc. Dyers Col.*, 86, 402
- Coates, E., Day, S., Provost, J. R., and Rigg, B. (1972), The measurement and assessment of colour differences for industrial use. 2. The accuracy of colour difference equations. *J. Soc. Dyers Col.*, 88, 69-75
- Davidson, H. R., and Friede, E., (1953) The size of acceptable color differences. *J. Opt. Soc. Amer.*, 43, 581-589
- Godlove, I. H. (1951), Improved color-difference formula, with applications to the perceptibility and acceptability of fadings. *J. Opt. Soc. Amer.* 41, 760-772
- Guilford, J. P. (1965), *Fundamental statistics in psychology and education*. McGraw-Hill, New York, p. 193
- Judd, D. B., and Wyszecki, G. (1963), *Color in business, science and industry*. John Wiley and Sons, New York, p. 267
- MacAdam, D. L. (1942), Visual sensitivities to color differences in daylight. *J. Opt. Soc. Amer.*, 32, 247-274
- MacAdam, D. L. (1964), Analytical approximations for color metric coefficients. III. Optimization of parameters in Friele's formulas. *J. Opt. Soc. Amer.*, 54, 1161-1165
- MacAdam, D. L. (1971), Geodesic chromaticity diagram based on variances of color matching by 14 normal observers. *Appl. Optics*, 10, 1-7
- Marshall, W. J., and Tough, D. (1968), Colour measurement and colour tolerance in relation to automation and instrumentation in textile dyeing. *J. Soc. Dyers Col.*, 84, 108
- McLaren, K. (1970), Colour passing—visual or instrumental? *J. Soc. Dyers Col.*, 86, 389
- Newhall, S.M., Nickerson, D., and Judd, D. B. (1943), Final report of the O. S. A. sub-committee on spacing of the Munsell colors. *J. Opt. Soc. Amer.* 33, 385-418
- Nickerson, D., and Stultz, K. F. (1944), Colour tolerance specification. *J. Opt. Soc. Amer.* 34, 550-570
- Thurner, K., and Walther, V., (1969). Untersuchungen zur Korrelation von Farbabstandsbewertungen auf visuellen Wege und über Farbdifferenzformeln. *Proc. 1st AIC-congress 'Color 69', Stockholm, 671-687, Farbe 18, 191-206*

The authors acknowledge the encouragement and financial assistance given by the Society of Dyers and Colourists.

DISCUSSION

Coates: It is perhaps relevant to draw attention to some of the reasons for the views expressed, and the type of approach made to the problem. We believe that the majority of colour-difference equations are based on data obtained under conditions which are not sufficiently close to the conditions of e.g. visual assessments in those industries concerned with surface colours. Either the field of view is too small, e.g. 2°, or the colour-differences employed are too large. In addition, the optimization of the equations to fit data through the whole of colour space we believe produces a situation with a fit to the surface-colour gamut which is worse than needs be. For these particular industries perhaps more success would be met by concentrating on the relevant region of colour space and the relevant conditions of assessment.

Hunt: I wonder whether it makes really sense to apply so much statistics,

on rather restricted material. Does not it make more sense to gather more data? The gamut is not adequately explored by far, and it seems to me that it is too soon to try to decide which is the best formula.

I know I have no right to speak, since I have not gathered any data myself on this subject.

Coates: Right; you have to do something however. More experimental data would be welcome and would allow more confidence to be placed on the statistical analysis. However, the analysis of any data of this type must be done statistically.

Jaekel: The trouble with more samples is that it is a lot of trouble. To get our "435" data, we had 560 patterns assessed, selected from about 1000 dyeings, and from planning to results took four years!

So we don't want to do a lot more of this. But the best is the enemy of the good, and with limited samples one can sort out the least suitable formulae for one's purpose quite considerably, and then concentrate on the bunch of better ones, and pick one of them for work trials.

Rigg: I agree completely.

**MULTIPLE LINEAR REGRESSION: A NEW
TECHNIQUE FOR IMPROVING COLOUR
DIFFERENCE FORMULAE**

K. McLAREN

*ICI organics division
Manchester M9 3DA England*

In 1963 the Council of the Society of Dyers and Colourists formed a Colour Measurement Committee, one of whose remits was to recommend the most suitable colour difference formula so as to prevent a situation developing whereby different organisations within the textile and allied industries would adopt different colour difference formulae. In 1970 sufficient evidence had been acquired (McLaren, 1970a) to show that formulae of the Adams-Nickerson type predicted the majority decisions of professional colourists engaged in visually appraising the acceptability of a commercial match better than any other, particularly the Official Recommendation of the CIE. As the Adams Chromatic Value formula had been most widely used in the UK, it was the obvious choice (Colour Measurement Committee, 1970); since then much additional data has been obtained which fortunately confirms the correctness of the committee's decision (Jaekel, 1971). It was known, however, that the colour space on which this formula was based, conveniently termed "ANLAB" (McLaren, 1970b) was far from being perfectly uniform and whilst it had been shown to be uniform enough to make instrumental shade passing a viable method (McLaren, 1969, 1970a; Jaekel, 1971), it was obvious that much greater reliability would be obtained if allowance could be made for these residual distortions: some random measurements had shown, for example, that if a difference of 1 AN (42) units was an acceptable limit for a grey, the limit for a saturated red could be as high as 3 units.

The Davidson and Friede data (1953) which were an important part of the evidence on which the committee's decision was based, consisted of the percentage acceptability and colorimetric data of 286 sample/standard pairs for 19 colour centres which were treated as a single population originally (McLaren, 1969, 1970a). It was considered that statistical analysis of each of the 19 sets

might give a reliable estimate as to how the pass/fail tolerance varied at each of these colour centres and thus provide a basis for data collection: it was even remotely possible that the variations might be systematic, so that the data could be extended by interpolation to cover a large proportion of the whole of surface-colour space. This further analysis was prompted by the availability of an extensive suite of statistical analysis programs (STATPACK) available to users of IBM's "Call 360" computer time-sharing service.

THE ESTIMATION OF THE PASS/FAIL LEVEL OF ΔE

The pass/fail level of ΔE (ΔE_{50}) is that value most likely to be accepted on 50% of the occasions the sample/standard pair was visually assessed. This is best estimated from a line drawn through the %A/ ΔE points which minimises the sum of the squares of the distance of each point from the line in the %A direction. This is termed a regression of %A on ΔE , the former being termed the dependent, the latter the independent variable though these terms are misleading because if one wanted to estimate the %A value most likely to yield a given value of ΔE from the same data, ΔE would be termed the dependent variable.

In the original study ΔE_{50} values were estimated from linear regressions of %A on ΔE (McLaren, 1969, 1970a), but since then it has been shown that curvilinear regressions fit the data better (McLaren, 1971a). Polynomial curves up to the sixth degree were, therefore, fitted to the %A/ ΔE values and the results for Set A are given in Table I (Goodness-of-fit is the percentage of the total sum-of-squares attributable to the regression; the *F* value is the variance ratio, marked* if significant at the 5%, ** if significant at the 1% and *** at the 0.1% level at the appropriate degrees of freedom).

Table I. Goodness of fit of % Acceptance versus ΔE curves with increasing polynomial degree for one set of data.

Variables	Polynomial Degree	Goodness-of-fit (%)	F Value
Set A %A/ ΔE	1	61.3	15.8**
	2	82.5	21.3***
	3	82.9	12.9**
	4	82.9	8.5**
	5	82.9	5.8*
	6	83.2	4.1

Analysis of variance showed that whilst the improvement in goodness-of-fit of a second degree (quadratic) polynomial regression over a linear regression was significant at the 1% level, the subsequent improvements were not signif-

icant even at the 25% level. This analysis was applied to the other 18 sets and showed the same pattern except in Sets *B* and *T* where no significant improvement over a linear regression was observed. Table II gives the significance of the fit of the best curve for each of the 19 sets together with the pass/fail levels, ΔE_{50} .

Table II. Significance of fit and pass/fail level ΔE_{50} for 19 data sets.

	Set	Significance	$\Delta E_{50}AN$ (42) units
	<i>A</i>	0.1	1.40
	<i>B</i>	1.0	1.50
	<i>D</i>	0.1	1.32
	<i>E</i>	20.0	1.56
	<i>F</i>	1.0	0.57
	<i>G</i>	0.1	0.88
	<i>H</i>	0.1	2.10
	<i>I</i>	0.1	1.07
	<i>J</i>	1.0	1.62
	<i>K</i>	0.1	1.11
	<i>L</i>	0.1	1.18
	<i>M</i>	10.0	1.38
	<i>N</i>	0.1	1.91
	<i>O</i>	0.1	1.37
	<i>P</i>	5.0	1.87
	<i>Q</i>	0.1	2.27
	<i>R</i>	0.1	1.85
	<i>S</i>	0.1	1.02
	<i>T</i>	5.0	1.60

The ΔE_{50} values varied from 0.57 in Set *F* (a dark grey) to 2.27 in Set *Q* (a light green) and it is obvious that if these values had been used to obtain the instrumental pass/fail decision instead of the single value of 1.3 calculated from the overall regression of %*A* on ΔE (McLaren, 1969, 1970*a*), the number of wrong decisions would be less than the 59 found in the latter case; the actual number is 41.

Determination of the appropriate pass/fail level for every colour centre is, however, rather impracticable, but the wide range of ΔE_{50} values found (4:1) suggested that it might be worthwhile to look for a systematic variation throughout colour space; fitting curves with ΔE_{50} as the dependent variable and the axes of ANLAB colour space, *L*, *A* and *B* in turn as the independent variable, seemed the simplest method of detecting this.

The results obtained are given in Table III. The effect of omitting the four ΔE_{50} values of lowest reliability (Sets *E*, *M*, *P* and *T*) was quite insignificant; they were, therefore, included in this and all subsequent analyses.

Table III. Survey of efforts to describe ΔE_{50} vs L , A , and B by curvilinear regression.

Variables	Curve	Goodness-of-fit (%)	F Value
$\Delta E_{50}/L$	Linear	29.5	7.1*
	Quadratic	30.4	3.5
	Cubic	30.4	2.2
$\Delta E_{50}/A$	Linear	13.6	2.7
	Quadratic	25.0	2.7
	Cubic	25.0	1.7
$\Delta E_{50}/B$	Linear	15.3	3.1
	Quadratic	20.0	2.0
	Cubic	20.0	1.3

Table III shows that there is a significant linear relationship between ΔE_{50} and L and this is given by the equation:

$$\Delta E_{50} = 0.78 + 0.0159L$$

This line is shown in Fig. 1 with the 19 individual points (X) and the mean point $\Delta E_{50}/L$ ($\circ$). However, the equation still leaves about 70% of the variance of ΔE_{50} unexplained and although this could not be accounted for by any relationship between ΔE_{50} and A or B it was possible that some function of A and B might show a relationship with ΔE_{50} . The first functions to be tried were the polar co-ordinates of the ANLAB uniform chromaticity scale diagram

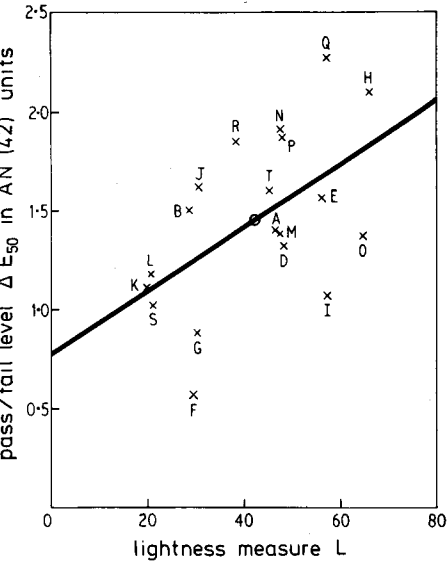


Fig. 1. The pass/fail colour difference, ΔE_{50} , plotted against the lightness measure, L . Letters indicate differently coloured data sets. The straight line is the best fitting curvilinear regression line. $\circ$ is gravity centre.

(McLaren, 1971*b*). Whilst the Cartesian co-ordinates A and B are ideal for plotting purposes, the visualisation of the colour of known AB values is facilitated by conversion to the polar co-ordinates C and T using the following relationships:

$$C = (A^2 + B^2)^{0.5} \quad T = \tan^{-1} (B/A)$$

C is thus the distance out from the neutral axis and, therefore, a measure of saturation; T is the hue angle starting with red at 0° followed by yellow at 90° , green at 180° and blue at 270° ; L , C and T , therefore, represent the variables of visual perception. The relationships with ΔE_{50} are given in Table IV and the points and quadratic curves are given in Figs 2 and 3. (The improvement, quadratic over linear, in the case of ΔE_{50} and C , however, was not significant at the 10% level).

Table IV. Survey of efforts to describe ΔE_{50} vs C (measure of saturation) and T (measure of hue) by curvilinear regression.

Variables	Curve	Goodness-of-fit (%)	F Value
$\Delta E_{50}/C$	Linear	39.3	11.1**
	Quadratic	45.0	6.6**
	Cubic	45.8	4.2*
$\Delta E_{50}/T$	Linear	1.3	0.2
	Quadratic	36.6	4.6*
	Cubic	36.8	2.9

It will be noted that the $A+$ radius has two T values and therefore the curve in Fig. 3 should give the same ΔE_{50} values at $T = 0$ and $T = 360$: it would be possible to fit such a curve but initially it did not seem necessary to do so.

MULTIPLE LINEAR REGRESSION

As the pass/fail level has been found to depend on three variables L , C and T the most reliable estimate of ΔE_{50} in any part of colour space bounded by the 19 colour centres will be given by a multiple linear regression equation. This was computed by the STATPACK stepwise regression program which introduces each independent variable into the equation in the order in which it accounts for the residual variation of ΔE_{50} . The relationship between ΔE_{50} and both C and T is, however, given by a quadratic equation and, therefore, the independent variables would be L , C , C^2 , T and T^2 .

Two pairs of these independent variables, C and C^2 ; T and T^2 , possess high correlation coefficients (over +0.9) and to reduce the possibility of this affecting the reliability of the multiple regression equation, the differences between each

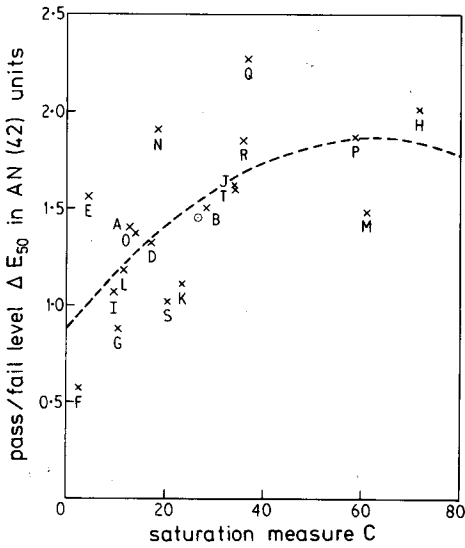


Fig. 2. The pass/fail colour difference, ΔE_{50} , plotted against the saturation measure, C . The Curve is the best fitting curvilinear regression line. $\odot$ is gravity centre.

value and their mean (e.g. $C - \bar{C}$) were taken as the values of the independent variables as this reduced the correlation coefficients to +0.7 and +0.06 respectively. To increase the accuracy of the mathematical treatment, the differences were further divided by the standard deviation of each independent variable.

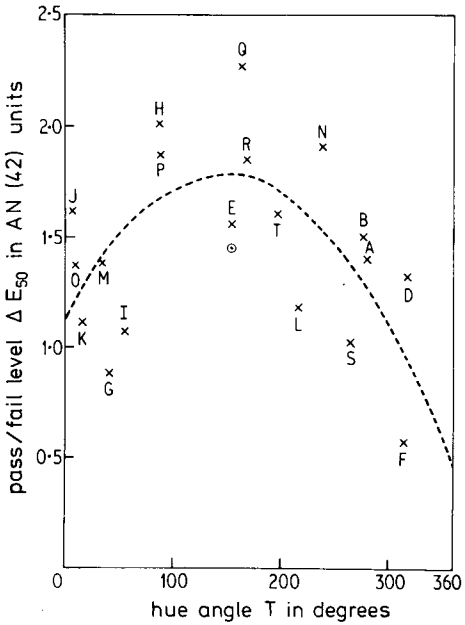


Fig. 3. The pass/fail colour difference, ΔE_{50} , plotted against the hue angle, T . The curve is the best fitting curvilinear regression line. $\odot$ is gravity centre.

The order in which the variables were introduced and the level of the significance of its contribution is given in Table V together with the cumulative goodness-of-fit.

Table V. Contribution to goodness of fit of subsequently introduced variables

Variable	Level of Significance (%)	Cumulative Goodness-of-fit (%)
<i>C</i>	0.5	39.4
<i>T</i> ²	2.5	57.9
<i>L</i>	10.0	66.9
<i>C</i> ²	1.0	79.7
<i>T</i>	>20.0	79.9

T was not entered into the regression because of the insignificance of its contribution and the regression equation obtained as follows:

$$\Delta E_{50} = 1.792 + 0.316 \left\{ \frac{C-26}{19} \right\} - 0.202 \left\{ \frac{T-154}{109} \right\}^2$$

$$+ 0.193 \left\{ \frac{L-42}{14} \right\} - 0.155 \left\{ \frac{C-26}{19} \right\}^2$$

Using this equation to predict the pass/fail level for each of the 19 sets resulted in 46 wrong decisions. The various numbers of wrong decisions are summarised in Table VI.

Table VI. Number of wrong decisions with various methods of assessing colour differences.

Method	No. of Wrong Decisions
Visual assessment—one observer	49
Instrumental—single ΔE_{50} value from ANLAB 42 formula	59
19 ΔE_{50} values from multiple regression equation	46
19 ΔE_{50} values from set data	41

Predicting the pass/fail level from the multiple regression equation is thus an effective method of reducing the number of wrong decisions had they been made instrumentally and fewer would have been made than would have occurred had any one colourist made the decisions on the basis of visual appraisal. The method could also be extended to other colour centres, but the particular visual pass/fail level would have no general validity nor would it even have applied to all the production of the firm involved (Alexander Smith, Inc.) as

this level applied only to "yarn to be used in a solid background color" (Davidson and Friede, 1953). However, it is possible to use the equation in any situation to obtain better agreement with visual assessments.

GENERAL APPLICABILITY OF THE MULTIPLE REGRESSION EQUATION

For any point in colour space within the gamut bounded by the Davidson and Friede standards the pass/fail borderline colour difference, ΔE_{50} , can be predicted; their values vary up to a maximum of 2.3 AN (42) units. All of these differences are visually equal (strictly speaking, they are equally acceptable but whilst acceptability and perceptibility may not always be identical, in this case they undoubtedly were).

At any point in ANLAB space, $L' C' T'$, the predicted pass/fail value, ΔE_{50} , will be n times that at a reference point L, C, T . Because the range of colour differences of industrial importance is only comparatively small, it is reasonable to assume that *any* colour difference at $L' C' T'$ will be numerically n times that of a visually equal colour difference at L, C, T ; one can, therefore, make *visually* equal differences *numerically* equal by dividing all $L' C' T'$ differences by n . The most appropriate reference point, L, C, T , is one giving ΔE_{50} equal to the mean of the 19 values which were originally calculated, i.e. 1.452, as this will tend to keep the average size of unit the same; more than one point will give this predicted value, of course, because many different L, C and T values will give $\Delta E_{50} = 1.452$.

Any colour difference in AN (42) units can therefore be corrected using the multiple regression equation by multiplying it by the factor

$$\frac{1.452}{1.792 + 0.193 \left\{ \frac{L-42}{14} \right\} + 0.316 \left\{ \frac{C-26}{19} \right\} - 0.155 \left\{ \frac{C-26}{19} \right\}^2 - 0.202 \left\{ \frac{T-154}{109} \right\}^2}$$

appropriately termed a multiple regression factor (MRF).

The means of the 250 Davidson and Friede ΔE values are 2.11 AN(42) units and 2.13 AN(42) $\times$ MRF units showing that the average size of unit is virtually unchanged.

Applying this multiple regression factor to the ΔE AN(42) values of the Davidson and Friede data improved the reliability of the formula when using either the correlation coefficient (McLaren, 1969, 1970a) or the goodness-of-fit (McLaren, 1971) as criteria, as shown in Table VII.

Application of the F test shows that the factor has improved the goodness-of-fit to a degree which is nearly significant at the 1% level.

It should be noted that the correlation coefficient for the Adams-Nickerson formula in Table VII is less than the -0.58 value found previously, (McLaren,

Table VII. Goodness of fit with and without introducing the Multiple Regression Factor MRF into the Adams-Nickerson formula.

Formula	Correlation Coefficient	Goodness-of-fit (Cubic Regression)
Adams-Nickerson	- 0.565	46.9%
Adams-Nickerson $\times$ MRF	- 0.648	60.1%

1969, 1970a). This is because the program used could only handle 250 %A/ ΔE values so every 8th pair was, therefore, omitted; reducing the number randomly invariably reduces the correlation coefficient.

A more realistic parameter of the reliability of colour difference formulae is the number of wrong decisions, i.e. the number of acceptable samples (%A > 50) rejected instrumentally plus the number of rejectable samples (%A < 50) accepted instrumentally. This requires an estimate to be made of ΔE_{50} and whilst this is best made from the regression of %A on ΔE , slight variations in value often has a marked effect on the number of wrong decisions. For appraising formulae it is, therefore, better to vary ΔE_{50} systematically and plot the number of wrong decisions for each value; this has been done for three formulae, viz. AN(42), AN(42) $\times$ MRF and NBS, the latter having been found to give the best agreement of all the established formulae against the Davidson and Friede data. These curves are shown in Fig. 4.

The use of the multiple regression factor has thus improved the reliability of the Adams-Nickerson formula by every criterion when it is applied to the Davidson and Friede data. This, however, only confirms the reliability of the

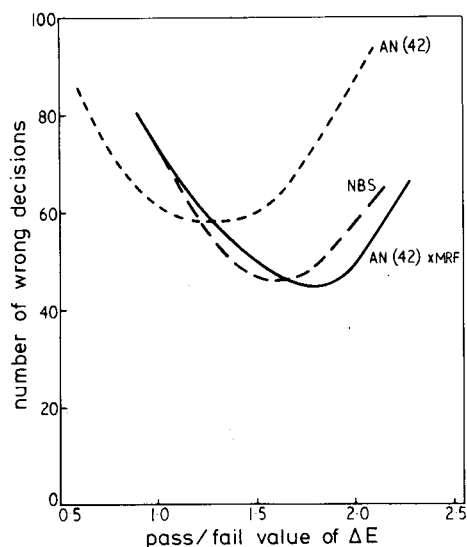


Fig. 4. The number of wrong decisions occurring at various pass/fail values for three formulae applied to the Davidson and Friede data.

multiple regression equation, not that of the modified formula because the test data is the same as that providing the factor; the equation on p. 302 must therefore never be used for routine colour difference measurement.

It would be too optimistic to expect that this particular factor would improve the reliability of the Adams-Nickerson formula when applied to other data but the technique of multiple linear regression is worthy of further investigation. Jaeckel (1971) has combined his own data with that of Davidson and Friede: if allowance could be made for the differences in visual tolerance of the colourists involved, multiple regression analysis should be applied to the total data (775 sample/standard pairs): if a multiple linear regression equation having a high goodness-of-fit could be derived for this data then the colour difference formula incorporating the corresponding MRF would stand a much better chance of being generally better than the Adams-Nickerson formula. Should the inclusion of other data such as that of Thurner and Walther (1969) result in another formula correlating best with visual assessments then the technique should be applied to this formula.

It should be noted that although the use of a multiple regression factor may result in an improved colour difference formula, it will not achieve this by defining a more uniform colour space. It would merely signify that perceptually two colours are closer or further apart than their distance in ANLAB(42) space implies. ANLAB(42)MRF space cannot exist because theoretically the distance between any two pairs of coloured samples *A* and *B* will depend on whether *A* or *B* is regarded as the standard. This is of no practical consequence for colour differences of industrial significance, but it is important in view of the fact that many believe that a uniform colour space cannot be Euclidean; if they are right, then multiple regression analysis may provide another method of getting round this limitation.

THE DISTORTIONS OF ANLAB SPACE.

From Figs 1 and 2 it can be deduced that pairs of samples exhibiting colour differences which are equally perceptible will be spaced further apart according to their distance upward and/or outward from the base of ANLAB space, the position of black. Whilst Fig. 2 suggests that at $C > 60$ they begin to close up again, this must not be assumed as extrapolation is not justified: this reinforces the plea of Schultze (1969) for more data relating to highly saturated colours. Fig. 3 shows that pairs of samples equally spaced visually are too widely spaced in the green region and too closely spaced in the red region.

ACKNOWLEDGEMENT

Whilst all the statistical analyses were carried out by the author using the Call/360 STATPACK programs, the correct interpretation of the results would

have been impossible without the assistance of Mr. J. D. Chamberlain; Dr. P. R. Bunkall suggested the particular method of extending the use of the multiple regression equation to any colour difference and Mr. J. T. Fallows wrote the program for calculating the number of wrong decisions; this help is gratefully acknowledged.

REFERENCES

- Colour Measurement Committee (1970), Recommended colour difference formula. *J. Soc. Dyers & Colourists* 86, 368
- Davidson, H. R., and Friede, E. (1953), The size of acceptable color differences. *J. Opt. Soc. Amer.* 43, 581-589
- Jaekel, S. M. (1971), The utility of colour difference formulae for match—acceptability decisions. This symposium.
- McLaren, K. (1969), The precision of textile colour matchers in relation to colour difference measurements. *Proc. 1st AIC congress 'Color 69', Stockholm*, 688-708 *Farbe* 18, 171-190
- McLaren, K. (1970a), Colour passing—Visual or instrumental? *J. Soc. Dyers & Colourists* 86, 389-
- McLaren, K. (1970b), The Adams-Nickerson colour difference formula. *J. Soc. Dyers & Colourists* 86, 354-
- McLaren, K. (1971a), The 'ANLAB' color difference formula for shade passing. Paper presented at the AATCC Golden Jubilee conference, Boston
- McLaren, K. (1971b), Tristimulus colorimetric specifications applicable in the textile industry. *J. Textile Inst.* 62, 453
- Schultze, W. (1969), Umfassender Vergleich von sieben Farbabstandsformeln. *Proc. 1st AIC congress 'Color 69', Stockholm*, 621-640 *Farbe* 18, 105-124
- Thurner, K., and Walther, V. (1969), Untersuchungen zur Korrelation von Farbabstandsbewertungen auf visuellem Wege und über Farbdifferenzformeln. *Proc. 1st AIC congress 'Color 69', Stockholm*, 671-687 *Farbe* 18, 191-206

DISCUSSION

Malkin: Have you considered interaction between L and C ?

McLaren: From Davidson and Friede's data, of course multiple regression analysis does do just this. We find that C is the most important variable, then T^2 , then L . This automatically takes care of the interaction, because, no matter whether L varies, or C , or T , you adjust ΔE accordingly. There cannot be any limitations from the available data that multiple regression cannot allow for.

Jaekel: If Peter Bunkall's suggestion works, that's the best justification. Nevertheless the argument may not totally correct, because linear regressions for different colours are rarely parallel, so relative equivalent ΔE 's for different severity levels differ.

McLaren: I accept your point. This is only a method of improving the reliability of colour difference formulae. When this is applied to your data I would be surprised if the superiority we found for the Davidson and Friede data would even partially be maintained; but, if you combined all your data, apply all

your formulae and then apply multiple regression analysis to the best of them, this may be a way of getting the best possible formula.

Jaeckel: Different sets of data differ: polynomial regression is less dramatically effective for some of our data than for the Davidson and Friede data.

McLaren: There is one thing that is particularly encouraging in Mr. Jaeckel's paper, with all his massive data and that is the consistent superiority of Adams type colour difference formulae. I cannot believe that this is a pure artefact.

Nickerson: Your multiple correlations interest me. But straight line multiple correlations, taking into consideration the factors of color specification—like hue, lightness and saturation, or hue, value and chroma, or the Adams values—do not necessarily give you the best answer.

For example, with multiple correlation carried to its extreme—i.e. curvilinear correlation—one may find that straight lines do not provide the best fit. We carried out such correlations some years ago when working with cotton standards and obtained greatly improved correlations of color with grade when the color factors were carried through to show the multiple curvilinear relationships. I think that a curvilinear type of investigation might lead to a better picture of the relationships that exist. At any rate, multiple correlation studies are very useful, and I would like to see them developed even further.

Vos: Your Fig. 3 puzzles me. How can you attribute different ΔE values to hue angles 0° and 360° ? Would not a sine function—giving a flat cross section through the hue cylinder—produce an equally good fit without this 0° discontinuity?

McLaren: Whilst I agree (p. 300) that the curve should give the same ΔE value at 0° and 360° and that it would be easy to fit such a curve, it seems pointless to do so on view of the extremely limited value of the derived MRF (p. 303) However, I would not, under any circumstances, fit a sine curve to the data because this predicts two points of inflection and there is absolutely no justification for believing that there is one at 300° .

26 INDUSTRIAL COLOR TOLERANCES BY XI-ETA FORMULA

FREDERICK T. SIMON

*Color Science Center, Clemson University
Clemson, SC 29631 USA*

INTRODUCTION

For many years industrial organizations have used color difference expressions to describe in objective terms how pairs of samples compare to each other. Notable among these formulas are the various interpretations and applications (Simon, 1947, Friele, 1965, Chickering, 1967) of the MacAdam (1942) experiments. At the time that the MacAdam metric was published much simpler means existed to calculate color differences, but the potential of using experimentally determined chromaticity discrimination data, rather than approximations, was more attractive to workers in the field, so it was adopted for many applications. MacAdam (1971) has recently published another metric that partly takes from the 1942 data (one observer) but has been expanded to include 14 observers and is given in an analytical form called *XI-ETA*. This paper discusses some data that has been obtained on a group of sample pairs of surface colors to compare several of the MacAdam-based metrics. Furthermore, a case is made for the better interpretation of color difference data with the use of descriptive words to define direction and kind of difference.

EXPERIMENTAL

A series of 30 woolen samples was chosen from the collection of swatches, Colorthek I, published by BASF. These were arranged in pairs so that the chromaticity difference, ΔC , between pairs was usually 3 to 7 (average 5) units. The colorants used to dye each pair were generally the same. By choice, the pairs were taken from many regions of color space and also several levels of lightness. All of the samples were measured against BaSO_4 on the Beckman

Table I. Comparison of MacAdam-based metrics

SPL.	Y	x	y	UNCORR. ΔC			CORR. ΔC			ΔL		
				SG	ξ, η	F1	SG	ξ, η	F2	SG	F1	F2
1260	3.34	0.2688	0.2754									
1254	3.62	0.2599	0.2614									
420	4.09	0.2373	0.1791									
414	4.13	0.2506	0.1906	16.1	9.4	11.5	9.1	5.4	8.6	0.3	0.3	0.1
610	5.25	0.2462	0.3035									
604	5.59	0.2331	0.2845	10.7	9.7	10.0	6.7	6.1	8.0	2.2	7.0	2.1
1043	7.29	0.4149	0.3317									
1037	7.92	0.4111	0.3417	8.8	6.4	8.4	6.2	4.5	7.4	3.2	7.4	2.6
527	8.03	0.2066	0.2033									
521	8.53	0.2097	0.1972	6.2	4.9	6.3	4.5	3.5	5.7	2.4	6.3	2.3
756	8.11	0.2848	0.4191									
762	9.50	0.2952	0.4388	5.6	7.7	7.3	4.2	5.7	6.6	6.5	16.2	5.9
252	8.15	0.5251	0.3079									
246	9.32	0.5332	0.3157	4.2	6.0	4.8	3.1	4.4	4.3	5.5	11.2	4.1
1008	8.37	0.4234	0.3827									
1002	9.71	0.4249	0.3944	5.0	5.3	5.9	3.8	4.0	5.4	6.1	15.4	5.7
180	10.89	0.5495	0.3457									
174	11.93	0.5472	0.3533	6.3	4.9	5.4	5.1	4.0	5.4	4.1	6.8	2.8
526	11.46	0.2159	0.2198									
520	11.81	0.2187	0.2141	5.8	4.3	5.3	4.6	3.4	5.5	1.4	3.2	1.4
1053	14.70	0.3688	0.3068									
1047	15.32	0.3771	0.3154	5.0	4.9	4.3	4.3	4.3	4.9	2.0	4.1	2.0
912	15.01	0.4800	0.4063									
906	18.72	0.4744	0.4178	7.9	6.4	9.9	7.5	6.0	11.3	11.4	22.6	10.9
279	15.73	0.3704	0.2624									
273	17.97	0.3829	0.2631	10.8	7.2	11.6	10.0	6.7	13.5	6.9	14.7	7.2
126	19.71	0.5348	0.3884									
132	20.52	0.5344	0.3846	1.6	2.0	1.9	1.6	2.0	2.4	2.2	3.9	2.1
405	19.80	0.2440	0.2247									
399	21.52	0.2544	0.2325	10.2	7.1	9.3	10.1	7.1	11.7	4.6	8.4	4.5
1112	19.92	0.3564	0.3171									
1106	21.15	0.3730	0.3270	10.8	8.7	10.2	10.6	8.5	13.0	3.3	6.5	3.5
278	22.78	0.3592	0.2740									
272	26.84	0.3675	0.2780	5.7	4.3	6.2	6.2	4.6	8.3	9.8	17.9	10.3
125	24.64	0.5256	0.3957									
131	25.84	0.5200	0.3870	3.2	5.2	3.7	3.4	5.6	5.2	2.9	4.7	2.8
1190	24.84	0.2985	0.3103									
1196	25.50	0.2962	0.3117	3.5	1.8	2.5	3.7	1.9	3.5	1.6	2.6	1.6
36	31.18	0.4480	0.4740									
42	35.30	0.4600	0.4686	10.2	6.4	9.7	12.3	7.8	14.6	8.3	14.2	9.2
1058	31.24	0.3418	0.3680									
1064	37.14	0.3518	0.3690	7.1	5.5	8.4	8.7	6.7	12.7	11.7	19.7	12.8
439	31.76	0.2858	0.2883									
433	33.38	0.2913	0.2898	5.1	3.4	4.6	6.0	4.0	7.0	3.3	5.3	3.4
1111	32.21	0.3455	0.3277									
1105	32.60	0.3595	0.3292	10.1	7.8	10.7	11.8	9.2	16.4	0.8	1.7	1.1
1027	33.75	0.3713	0.3311									
1021	35.18	0.3801	0.3310	6.3	4.8	6.7	7.7	5.7	10.4	2.8	4.7	3.1
829	33.95	0.3247	0.3686									

Table I. (Continued)

SPL.	Y	x	y	UNCORR. ΔC			CORR. ΔC			ΔL		
				SG	ε,η	F1	SG	ε,η	F2	SG	F1	F2
823	34.06	0.3265	0.3757	2.3	2.8	2.3	2.7	3.3	3.6	0.2	0.1	0.1
1213	37.06	0.3038	0.3301									
1219	39.31	0.3022	0.3322	2.7	1.5	2.1	3.5	1.9	3.3	4.2	6.1	4.2
998	37.69	0.3497	0.3644									
992	39.45	0.3624	0.3754	7.4	7.0	7.1	9.4	8.9	11.4	3.2	4.7	3.3
193	41.19	0.4169	0.3513									
199	47.38	0.3997	0.3436	8.2	8.6	11.2	11.2	11.9	18.6	10.5	14.3	10.1
979	44.46	0.3505	0.3622									
973	48.13	0.3639	0.3813	8.5	9.1	7.7	11.8	12.6	13.1	6.0	8.1	5.9
931	48.31	0.3835	0.3568									
925	50.43	0.3775	0.3669	7.6	6.6	8.9	10.7	9.3	15.6	3.3	3.7	2.7

Explanation of Table I

Uncorrected Delta C

SG Simon-Goodwin (Simon, 1949). A computer program was used instead of graphical computation.
εη MacAdam (1971).
F1 FMC-1 (Chickering, 1967).

Corrected Delta C

An analytical expression in terms of Munsell value, V, was used to 'correct' the ΔC calculated by both the S-G and XI-ETA formulas. This expression is:

$$H = (0.13V - 0.004V^2 + 0.003V^3)$$

and is adapted from the graphical method (Simon, 1949).

$$\begin{aligned} SG_{corr.} &= (\Delta C_{uncorr.}) * H \\ \epsilon\eta_{corr.} &= (\Delta C_{uncorr.}) * H \\ F-2 &= (\Delta C_{FMC-1}) * K1 \text{ (Chickering, 1969)} \end{aligned}$$

Delta L

$$\begin{aligned} SG &= (V_{standard} - V_{sample}) * 25.0 \text{ (Simon, 1949)} \\ F1 &\text{ (Chickering, 1967)} \\ F2 &\text{ (Chickering, 1969)} \\ V &= \text{Munsell Value} \end{aligned}$$

Color DB-G spectrophotometer and the data were recorded on punched paper tape and reduced on a digital computer.

Table I gives a summary of the calculated color differences based on four Mac Adam metrics, Simon-Goodwin, XI-ETA, FMC-1 and FMC-2. The data are arranged so that the sample in the pair is always lighter than its comparison standard, and they are given in ascending order of tristimulus, "Y".

It is not shown in the table that there is substantial, although not complete, agreement among the ΔC results calculated by the four metrics. However, the

ΔL data from FMC-1 do not agree with the other two. The data in Table I were then examined to find sets of ΔC computations where importantly different numbers were obtained by any one formula which really should agree with the other two. These examples were then judged visually by several observers to determine which formula ranked the difference correctly. For example, if visual tests rate pair 526-520 as having a smaller chromaticity difference than pair 1053-1047, this favors the *XI-ETA* formula over the other two.

No exterior visual testing was done but about 12 such pairs were examined and a tally showed that the *XI-ETA* formula was more consistent with visual estimates than either S-G or FMC computations.

It will be noted that each of the parameters, ΔC and ΔL , are treated separately so that ΔL can be weighted differently as recommended by Judd (1963). It has been the author's experience that colored materials of differing texture require a weighting factor for ΔL in the S-G formula of about 0.5 for a very glossy plastic to 2 or 3 for matte textiles, with factor for ΔC equal to one.

Furthermore, aside from the psychological effects upon perceived color of surround, size of object, and dividing line, there is a considerable range of tolerance for lightness that depends entirely upon the end use of the colored product. Acceptable lightness limits cover a much greater gamut from industry-to-industry than the chromaticity. To say it another way, in order to understand color difference data, it is imperative to separate the major components of color, chromaticity and lightness.

It is tempting to oversimplify color differences and only use numbers to describe the relation between pairs of samples. In point of fact, numbers alone are quite unrewarding in conveying ideas among people, unless, of course, a statistical study is being made. One of the reasons for widespread acceptance of the Hunter (1958) and Glasser (1958) scales is the simple interpretation of the direction of the difference data in terms of four primaries that are related to readings taken directly from instruments. Chickering (1969) developed the method for the FMC equation to resolve the chromaticity difference vector into its component yellow-blue and red-green axes. Although these principles are eminently useful, I prefer a somewhat more elegant approach to the problem based upon the idea of "unitary hues" described by Burnham (1963). Actually, the description of one blue as "bluer" than another when it is closer to the Hunter "+a" coordinate, does not make sense. If one accepts the limits described by Burnham for the four "unitary hues", it is easy to define regions of unitary and secondary hues which do not get into conflicting descriptions. Thus, pairs of blues are either redder or greener whereas violets or purples are bluer or redder, and so forth. This is a more understandable system that is shown in Fig. 1 and can be readily adapted to simple geometric computation.

Two other sets of useful words that convey additional information about the *kind* of color difference between samples are the terms "lighter-darker"

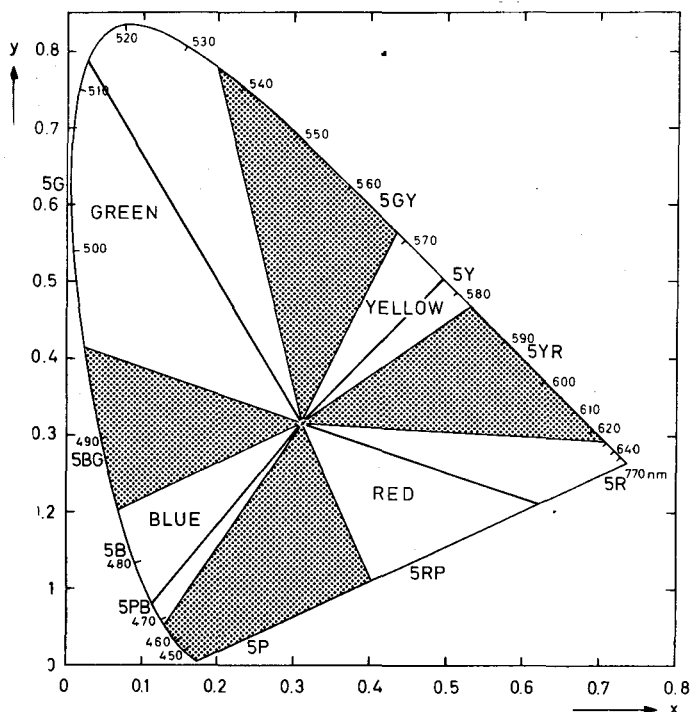


Fig. 1. Unitary hue areas in the CIE mixture diagram.

and "brighter-duller". Certainly no explanation is needed to attach significance to the direction of lightness differences. However, the suggested saturation adjectives are associated with the textile industry and have widespread acceptance among its members. On the other hand, very little explanation is needed to gain understanding by the uninitiated.

With all of the foregoing in hand, the color difference data calculated by *XI-ETA* for the same set of samples is given in Table II, but with word descriptors. The computed Munsell approximations (Simon, 1971) are also given with each sample to provide additional information about the samples in terms of a visually-oriented system.

CONCLUSION

This brief study has incorporated some of the principles that I believe are important to understand the relation between color difference expressions. 1) If numerical comparisons are to be studied, it is necessary that results be justified with spectrally similar pairs before metameric pairs are computed, 2) the components of the difference, chromaticity and lightness should be con-

sidered separately and not together as ΔE , 3) certain visual bench marks such as Munsell notations be used to determine the relevance of the scaling of the parameters wherever possible, and 4) word descriptions should be used with all numerical difference data to aid in interpretation.

Table II. Descriptive color difference data

SPL.	XI*	ETA*	MUNSELL	ΔC	ΔL	ΔE
1260	-8.0	-18.6	5.4PB	2.08/ 1.7		
1254	-10.0	-24.9	6.1PB	2.18/ 2.4	3.6 RED-BRI	2.5 LT 4.4
420	1.4	-57.7	2.4 P	2.34/ 7.8		
414	8.5	-51.4	3.2 P	2.35/ 7.2	5.4 RED-DUL	0.3 LT 5.4
610	-30.6	-12.5	7.5BG	2.68/ 2.2		
604	-35.3	-21.1	2.1 B	2.77/ 2.7	6.1 BLU-BRI	2.2 LT 6.5
1043	58.6	26.1	5.5 R	3.16/ 4.8		
1037	53.2	29.4	6.9 R	3.29/ 4.4	4.5 YEL-DUL	3.3 LT 5.5
527	-35.9	-54.4	5.2PB	3.31/ 7.0		
521	-31.2	-56.1	6.7PB	3.41/ 7.6	3.5 RED-DUL	2.4 LT 4.3
756	-31.4	30.1	1.8 G	3.33/ 4.7		
762	-29.7	37.5	0.6 G	3.59/ 5.3	5.7 RED-BRI	6.5 LT 8.6
252	118.1	36.4	4.1 R	3.34/10.8		
246	117.9	42.5	4.7 R	3.56/11.2	4.5 YEL-BRI	5.5 LT 7.1
1008	45.4	47.3	3.6YR	3.38/ 4.0		
1002	42.5	51.7	5.7YR	3.63/ 4.2	4.0 YEL-BRI	6.2 LT 7.3
180	111.9	62.3	7.0 R	3.83/11.1		
174	108.1	65.5	7.6 R	3.99/11.1	4.0 YEL-DUL	4.1 LT 5.7
526	-32.2	-46.8	4.0PB	3.92/ 6.6		
520	-28.2	-48.3	5.7PB	3.97/ 7.1	3.4 RED-DUL	1.4 LT 3.7
1053	43.5	8.4	3.8 R	4.38/ 5.0		
1047	45.0	13.0	5.7 R	4.47/ 5.2	4.3 YEL-BRI	2.0 LT 4.7
912	62.8	69.6	3.9YR	4.43/ 7.2		
906	57.0	72.1	5.6YR	4.88/ 7.5	6.0 YEL-DUL	11.4 LT 12.9
279	62.5	-9.1	5.0RP	4.52/ 8.7		
273	69.5	-7.2	5.7RP	4.79/ 9.9	6.7 YEL-BRI	6.9 LT 9.6
126	90.6	77.9	0.7YR	4.99/11.4		
132	91.7	76.2	0.3YR	5.08/11.6	2.0 BLU-DUL	2.2 LT 3.0
405	-10.3	-40.4	9.1PB	5.00/ 7.7		
399	-4.7	-36.0	0.1 P	5.19/ 7.3	7.1 RED-DUL	4.7 LT 8.5
1112	32.9	10.2	6.1 R	5.02/ 4.2		
1106	38.6	16.8	5.2 R	5.15/ 4.6	8.5 YEL-BRI	3.3 LT 9.2
278	50.8	-5.9	5.3RP	5.32/ 8.0		
272	54.1	-3.2	6.3RP	5.71/ 8.8	4.6 YEL-BRI	9.8 LT 10.9
125	84.5	78.2	1.4YR	5.51/11.6		
131	85.3	73.0	0.5YR	5.62/11.7	5.6 BLU-DUL	2.9 LT 6.3
1190	1.0	-1.7	4.0 P	5.52/ 0.4		
1196	-0.8	-1.6	3.7PB	5.59/ 0.3	1.9 GRN-DUL	1.6 LT 2.5
36	30.7	82.6	4.4 Y	6.10/ 9.1		
42	36.9	84.3	3.0 Y	6.43/ 9.8	7.8 RED-BRI	8.4 LT 11.4
1058	9.4	25.5	7.1 Y	6.10/ 2.6		
1064	14.3	27.7	3.8 Y	6.57/ 3.0	6.8 RED-BRI	11.8 LT 13.6
439	-0.4	-11.5	9.2PB	6.14/ 2.4		
433	2.7	-10.1	3.1 P	6.28/ 2.5	4.0 RED-DUL	3.3 LT 5.2
1111	23.3	12.3	6.6 R	6.18/ 3.1		

Table II. (continued)

SPL.	XI*	ETA*	MUNSELL		ΔC		ΔL		ΔE	
1105	30.6	15.2	6.1	R	6.21/ 4.2	9.2	BLU-BRI	0.8	LT	9.2
1027	36.3	18.0	6.1	R	6.31/ 5.1					
1021	40.7	19.9	6.0	R	6.42/ 5.8	5.8	BLU-BRI	2.8	LT	6.4
829	0.2	22.5	3.8	GY	6.32/ 2.9					
823	-0.7	25.1	4.2	GY	6.33/ 3.3	3.3	GRN-BRI	0.2	LT	3.3
1213	-1.4	6.0	2.9	GY	6.57/ 1.4					
1219	-2.9	6.4	4.6	GY	6.73/ 1.7	1.9	GRN-BRI	4.2	LT	4.6
998	14.5	25.8	2.9	Y	6.61/ 2.8					
992	18.0	32.0	2.5	Y	6.74/ 3.6	8.9	GRN-BRI	3.3	LT	9.5
193	52.6	34.2	8.1	R	6.87/ 7.9					
199	46.8	27.9	7.4	R	7.29/ 7.3	11.9	BLU-DUL	10.5	LT	15.9
979	15.6	25.2	1.7	Y	7.10/ 3.0					
973	17.1	34.2	3.7	Y	7.34/ 4.1	12.7	YEL-BRI	6.1	LT	14.0
931	34.2	29.7	1.4	YR	7.35/ 5.3					
925	28.1	32.1	5.3	YR	7.48/ 4.7	9.3	YEL-DUL	3.4	LT	9.9

Explanation of Table II.

ΔC According to MacAdam (1971) but corrected as in Table I

ΔL According to Simon (1949)

Munsell Approximations According to Simon (1971)

Hue Descriptors According to Burnham (1963)

REFERENCES

- Burnham, R. W., Hanes, R. M. and Bartleson, C. J. (1963), *Color: A Guide to Basic Facts and Concepts*. Wiley, New York
- Chickering, K. D. (1967), Optimization of the MacAdam Modified 1965 Friele Color Difference Formula. *J. Opt. Soc. Amer.* 57, 537-541
- Chickering, K. D. (1969), Perceptual Significance of the Differences between C.I.E. Tristimulus Values. *J. Opt. Soc. Amer.* 59, 986-990
- Friele, L. F. C. (1965), Further Analysis of Color Discrimination Data. *J. Opt. Soc. Amer.* 55, 1314-1319
- Glasser, L. G., McKinney, A. H., Reilly, C. D. and Schnelle, P. D. (1958), Cube-Root Color Coordinate System. *J. Opt. Soc. Amer.* 48, 736-740
- Hunter, R. S. (1958), Photoelectric Color Difference Meter. *J. Opt. Soc. Amer.* 48, 985-995
- Judd, D. B., and Wyszecki, G. (1963), *Color in Business Science and Industry*. Wiley, New York
- MacAdam, D. L. (1942), Visual Sensitivities to Color Differences in Daylight. *J. Opt. Soc. Amer.* 32, 247-274
- MacAdam, D. L. (1971), Geodesic Chromaticity Diagram Based on Variances of Color Matching by 14 Normal Observers. *Applied Optics* 10, 1-7
- Simon, F. T., and Alexander, J. F. (1971), A Scientific Method for Cataloguing Colored Samples. To be published
- Simon, F. T. and Goodwin, W. J. (1949), Rapid Graphical Computation of Small Color Differences. *Amer. Dyestuff Reporter* 47, 105-112

27 FMC-METRICS: WHAT NEXT?

L. F. C. FRIELE

*Fibre Research Institute TNO
Delft, Netherlands.*

~~DEVELOPMENT OF FMC-METRICS~~ ~~COMPARISON OF LARGE AND SMALL COLOUR DIFFERENCE SCALINGS.~~

At the international colour conference in Düsseldorf I presented a paper (Friele, 1961) in which the Müller stage or zone theory was applied to analyse the threshold ellipsoids of Brown (1957) and Brown-MacAdam (1949).

The zonal processing of the cone stimuli resulted in an orthogonal coordinate system for a lightness and two chromatic signals. The threshold formulation was obtained by applying Weber's law to the individual cone responses, introducing however some domination principles.

The following line element resulted:

$$(\Delta E)^2 = (\Delta L)^2 + (\Delta C_{r-g})^2 + (\Delta C_{y-b})^2$$
$$\Delta L = \frac{1}{2\alpha} \left(\frac{\Delta R}{R} + \frac{\Delta G}{G} \right)$$
(1)

$$\Delta C_{r-g} = \frac{1}{\beta} \left(\frac{\Delta R}{R} - \frac{\Delta G}{G} \right)$$

$$\Delta C_{y-b} = \frac{1}{\gamma} \left(\frac{\Delta R}{2R} + \frac{\Delta G}{2G} - \frac{\Delta B}{B} \right)$$

The domination principles were expressed in the definition of α , β and γ . R , G and B are the cone stimuli as calculated from X , Y and Z by means of a linear transformation. Obviously the directions of the lightness and the two chromatic processes in RGB space are obtained by successively putting the variation Δ of two of them equal to zero.

This line element was studied and amended by MacAdam in a series of

papers (see Billmeyer, 1969) from which he concluded this line of approach to be a very promising one. The Eqs (1) had however the drawback to fail (Friele, 1966) for boundary conditions as e.g. for $G \rightarrow 0$ also $\Delta G \rightarrow 0$. For this reason and others a modification was undertaken. Modification of the threshold formulation could however change the predicted directions of the lightness and chromatic processes in RGB space.

In my second proposal (Friele, 1965) I therefore transformed the RGB space first to an orthogonal coordinate system for the lightness and chromatic processes and then formulated the threshold definition.

The formulation was based again on the Brown and Brown-MacAdam ellipsoids, but furthermore the MacAdam (1942) ellipses, the Brown (1951) ellipsoids for a range of luminance levels and the Davidson-Friede data (1953) on commercial tolerances were added to the analysis.

This improved version was again amended by MacAdam and then optimized by Chickering (1967). In this form it is known as the FMC-1 colour difference formula. During its development it had happily passed birth and childhood and had grown up to a mature individual, mainly due to the lasting activity of its guardian MacAdam.

The FM-2 formula will be discussed later.

EXPECTED FURTHER DEVELOPMENT

The concept of zonal processing of the cone stimuli has proved to be a powerful tool in the analysis of threshold discrimination. The Chickering optimization is based on the MacAdam modified 1965 Friele metric in its version for high luminance levels and is valid for the 1942 MacAdam ellipses. The 1965 proposal was however also presented in a version for low luminance levels by replacing the Weber law by the de Vries-Rose law.

In Fig. 1 luminance discrimination is presented according to the Brown-MacAdam (1949) and Brown (1951) ellipsoids for a selected number of colour centres. Luminance discrimination was calculated as the intersection of the ellipsoid in the direction for which the variation of the two chromatic processes is zero, and is expressed as ΔY versus Y . The upper part suggests a discrimination of the Weber type; the lower part gradually changes into the de Vries-Rose type. The same transition is observed for the discrimination in the chromatic signals.

In Fig. 2 luminance discrimination is plotted as calculated by means of the 1965 metric. It falls apart in two branches, one of the Weber and one of the de Vries-Rose type, without a transition zone which links the two branches in a smooth way. This is caused by the fact that the 1965 proposal handled high and low luminance levels with separate threshold formulations. Moreover the

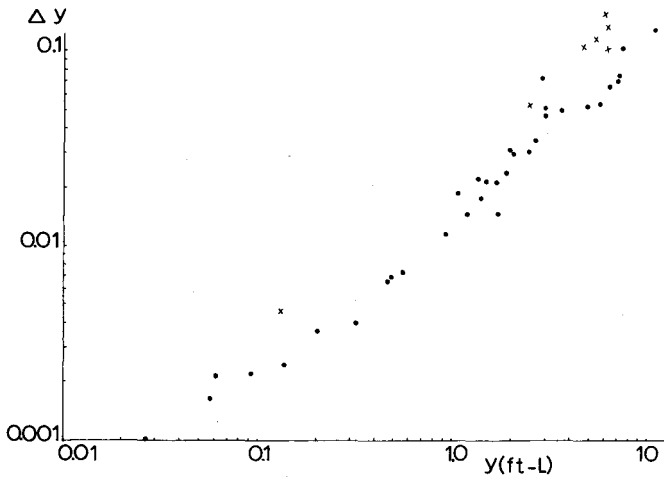


Fig. 1 Luminance discrimination for a number of colour centres selected from the Brown-MacAdam (1949) and Brown (1951) ellipsoids · × saturated reds.

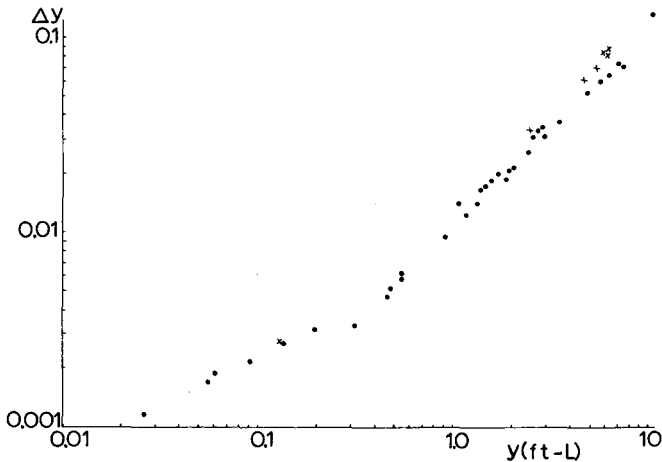


Fig. 2. Luminance discrimination for the same colour centres as in Fig. 1 according to the 1965 Friele metric.

experimental results (Fig. 1) show a detail for saturated reds, which is less pronounced in the calculated thresholds (Fig. 2).

Now the development of the 1965 metric was based on the qualitative Müller zone theory. The idea of zonal processing has got a firm experimental confirmation more recently. Furthermore the nature of low and high luminance discrimination principles and their interconnection has been cleared up by the work of Bouman and coworkers. These principles were applied by Wal-

raven in a zone-fluctuation model of colour discrimination, followed by a line element suggested by Vos and Walraven (1972). This line element was tailored to fit wavelength discrimination data. Some preliminary calculations however revealed it to be less successful in the description of threshold ellipsoids. It is intended to follow the line of approach in a joint program of the Institute for Perception TNO and the Fibre Research Institute TNO to develop a line element which explains threshold ellipsoids and wavelength discrimination data equally well by means of the zone-fluctuation model. The new line element may then be the successor of the 1965 Friele metric and of FMC-1 and represent up to date physiological knowledge.

COMPARISON OF LARGE AND SMALL COLOUR DIFFERENCE SCALINGS

The Müller zone theory has been also applied to analyse (Friele 1969) the Munsell system which is supposed to represent scaling principles for large colour differences. This analysis revealed scaling principles which are basically different from those underlying threshold discrimination.

This can be demonstrated most clearly by Fig. 3 and Fig. 4 which illustrate the discrimination principles for the yellow-blue chromatic response. Assuming zero or a very small contribution of the blue cone response to luminance, the direction of the yellow-blue response is parallel to the B axis in RGB space. Discrimination in the yellow-blue response is therefore measured by ΔB for equiluminous colours.

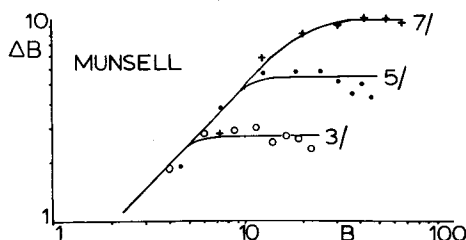


Fig. 3. The discrimination principle in the yellow-blue chromatic response for large colour differences (Munsell Δ Chroma = 2).

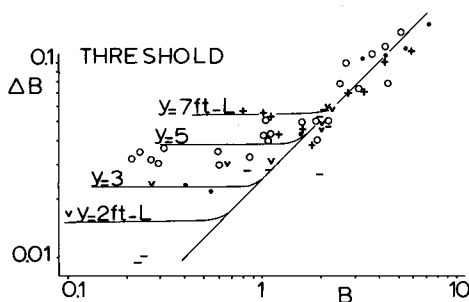


Fig. 4. The discrimination principle in the yellow-blue chromatic response for threshold differences: Brown and Brown—MacAdam data. + $Y = 6-8 \text{ ft} - L$ o $Y = 4-6 \text{ ft} - L$. $Y = 2.5-4 \text{ ft} - L$ v $Y = 1.5-2.5 \text{ ft} - L$ — $Y = < 1.5 \text{ ft} - L$

In Fig. 3 ΔB is related to B , for differences of 2 chroma units (Munsell system); in Fig. 4 the same plot is used, however, representing threshold differences. The interpretation of these figures is the following. (Friele, 1971).

For threshold differences the limitation in discrimination is caused by fluctuation or noise in the signal. The noise being proportional to the receptor signal (Weber behaviour) the noise in the yellow-blue signal is the resultant from the two noise levels fed into the opponent response $Y-B$, that is the noise level ΔY in the Y signal and the noise level ΔB in the B signal. Obviously the effective noise level will be ΔB for $B \gg Y$ and ΔY for $Y \gg B$. This is indeed demonstrated by Fig. 4, where one recognises a proportionality of ΔB and B for relatively high B value. However, for relatively low B value the threshold ΔB becomes independent on B but proportional to Y .

For large colour differences (Fig. 3) a reversed picture is obtained, ΔB being proportional to B for relatively low B values (Weber's law) but becoming independent on B and dependent on Y for relatively high B values.

The spacing of the horizontal branches is such as to be expected for a cube root response for Y .

Fig. 3 indicates that the opponent yellow-blue response discrimination is according to the Y or to the B input dependent on which of the two alternatives gives the best resolution. So the principle of opponent response yields a better resolution in visual discrimination than would have been possible without zonal processing.

The price to be paid for this achievement is the fact that for threshold differences the visual system is condemned to use the worst of the two possibilities available and this is due to noise summation. The threshold metric is then only a borderline metric.

INDUSTRIAL TOLERANCES

Industrial tolerances are, generally spoken, equivalent to colour differences somewhat larger than just perceptible or threshold differences, but distinctly smaller than steps of 2 chroma units. For sake of comparison: a difference of 2 chroma units, is in a way equivalent tot 28 MacAdam or threshold units (Friele, 1969). Industrial tolerances are mainly well within 5 MacAdam units. So for industrial purposes we are mainly interested in the area just beyond the borderline and the question is accute which metric is suited to measure colour differences of commercial importance.

The discussions on this point are hot and partly very confusing. However, when we keep our heads cool it is possible to discern some distinct contours through the smoke of the battle.

All colour difference formulas used for the evaluation of tolerances are based on transformations or distortions of the CIE standard colorimetric

system, which is a visually non-uniform system without doubt.

Visual data on tolerance should be obtained in such a way to enable the construction of an ellipsoid of equal visual tolerance around a standard colour. Analysis of the scaling properties around the standard colour and interrelation of the results for many colour centres scattered over the colour gamut then enables the construction of a metric. Knowledge of up to date physiological evidence will facilitate the interpretation of the data and the construction of the frame work of the formula. This is a cumbersome procedure, which however cannot be avoided.

Now the derivation of the 1961 and 1965 Friele metrics was indeed based on a time-consuming analysis of ellipsoid axes directions and lengths, followed by the formulation in analytical form, which gave a fair representation of the discrimination rules observed. Analysis of the Davidson-Friede data led to the conclusion that the scaling principles for commercial tolerances can be represented by the same type of metric as obtained for threshold ellipsoids. The numerical values of the parameters however have to be adapted. Thus the discrimination principles for colour differences of commercial importance are very much alike those for threshold discrimination and therefore quite different from scaling principles for large colour differences.

It makes no sense to start straightaway a comparison of colour difference formulas with visual data in a statistical way for observations all over the colour domains unless a perfect correlation would turn out; this is however impossible in view of the statistical character of the visual data themselves. In the overall correlation unfavourable properties of the formula in certain regions could well be masked by favourable scaling in other regions and one is always interested in the reliability per colour.

The spatial scaling properties of available colour difference formulas can be very different indeed as was demonstrated most clearly by Schultze (1969) at the Colour meeting in Stockholm. From a theoretical and experimental verification of the combination of luminance and saturation differences Schultze and Gall (1969) concluded that only the Simon-Goodwin system and the FMC-2 and DIN formulas behaved well in this respect. FMC-2 is a modification of FMC-1; FMC-1 and FMC-2 do not differ in this respect.

Regarding chromaticness scaling Davidson and Friede (1953) concluded that the Mac-Adam ellipses correlate better with their visual data on commercial tolerances than do the other methods examined by them. This conclusion confirms the above statement about similarity of threshold and tolerance metrics.

FMC-1 and FMC-2 are both based on the MacAdam ellipses and we have now to discuss the amended FMC-1 formula which is known as FMC-2 (Chickering 1971) and which is one of the four formulas recommended for further field trials by CIE. The idea to amend FMC-1 originated from a comparison of FMC-1 and the Simon-Goodwin system, both being based on MacAdam

ellipses. In the Simon-Goodwin system the size of these chromaticity ellipses is varied dependent on luminance in the same way as chroma contours shrink or swell in the Munsell system. For the lightness scale the Simon-Goodwin system utilizes the Munsell value scale. As in FMC-1 the chromaticity ellipses are independent of luminance and furthermore the lightness scale is based on Weber's law, two factors k_1 and k_2 were added to the FMC-1 metric to adapt it to the Simon-Goodwin system.

The FMC-2 formula is therefore a hybrid of threshold ellipsoids and the Munsell system. The basic differences between threshold and Munsell scalings make this hybridization suspect. Moreover Davidson and Friede (1953) concluded that their data do not consistently demonstrate a dependence of the size of ellipsoids on apparent luminous reflectance. Their Fig. 10 is an argument against the utilization of the Munsell value scale.

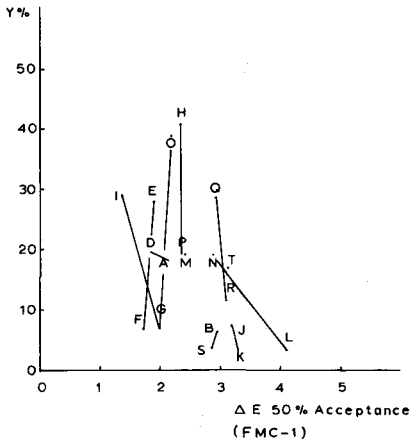


Fig. 5. ΔE for the 50% acceptance level, calculated from the *D-F* data by means of FMC-1. FMC-1 was adapted to the *D-F* data by changing two parameter values: $l = 0.086$ and $f = 1.2$.

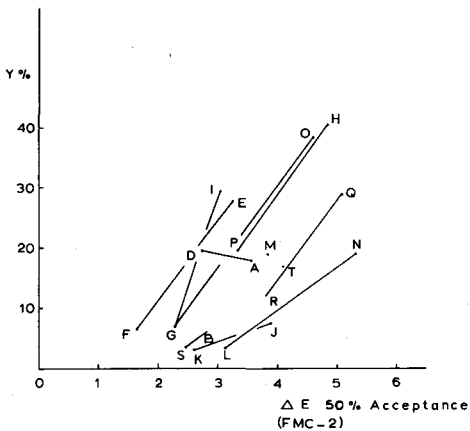


Fig. 6. ΔE for the 50% acceptance level, calculated from the *D-F* data by means of FMC-2.

In Fig. 5 the colour difference ΔE for the 50% acceptance level per colour centre as calculated from the Davidson-Friede data by means of regression analysis is plotted against luminance. ΔE was calculated by means of the FMC-1 formula in which however the parameter values for l and f were adapted to fit the $D-F$ data. Points representing colour centres of nearly equal chromaticity but of different luminance are connected by lines. It is obvious that there is no indication for a marked dependency of the ellipsoid size on luminance. This confirms the conclusion by Davidson and Friede mentioned above.

In Fig. 6 the same plot is made however using FMC-2. This figure proves most clearly that the amendment of FMC-1 is in error. This is a warning against hybridization of aspects of colour metrics of very different character. Modifications of existing formulas to adapt them to special data should be done with utmost care. As an example, it is often assumed that lightness and chromaticness contributions can be modified independently, e.g. that a logarithmic lightness scale can be replaced by a cube root scale without consequences for the chromaticness scaling. Lightness and chromaticness scaling are however interrelated. Different response functions for lightness are accompanied by differences in the chromaticness response functions as well and vice versa.

Although the $D-F$ data are of high quality the question arises whether they are the only tolerance data to be relied upon. The information in this field is increasing rapidly in recent years. The Thurner-Walther (1969) data are worth mentioning. They were analysed by me (unpublished results).

The conclusions drawn from this analysis are in many respects parallel to those for the $D-F$ data. The size of chromaticity ellipses of equal tolerance is not distinctly dependent on luminance. There is however a clockwise rotation of tolerance ellipses as compared to the MacAdam ellipses, as was observed by Davidson-Friede (1953) for their data. A separate analysis of luminance scaling proved it to be of the Weber type and certainly not according to the Munsell value function. The MacAdam modified 1965 Friele metric is applicable.

Kuehni (1970) discussed the relationship between acceptability and calculated colour differences by means of the $D-F$ and $T-W$ data. His figures give the impression that the $T-W$ data are less reliable, in accordance with my own conclusion. I found a considerable increase in spread for dark colours as compared to lighter colours, and advise to exclude the data for colours with a luminance value below 10% from further analysis of the $T-W$ data or at least to handle them separately.

The FMC-2 metric being a failure it should be abandoned in future evaluations of colour difference formulas. The structural properties of the MacAdam modified 1965 Friele metric seem to be a fair approximation of the visual scaling principles in rating colour differences of commercial importance.

This cannot be said for most of the other existing formulas. The Chickering

optimization (FMC-1) is valid for MacAdam ellipses. The optimized parameter values should however not be applied for tolerance evaluation without verification.

Chickering obtained the following parameter values for the MacAdam ellipses:

$$\begin{aligned}\alpha &= 0.00416 & P &= 0.724X + 0.382Y - 0.098 Z \\ \beta &= 0.0176 & Q &= -0.480X + 1.370Y + 0.1276Z \\ p &= 0.4489 & S &= 0.686 Z \\ f &= 0 \\ N &= 2.73 \\ l &= 0.279\end{aligned}$$

For the *D-F* data the correlation coefficient for linear regression of acceptance and ΔE for the acceptance range 95 – 5% only was calculated, using the FMC-1 formula. Then the parameter values l and f were varied. Optimal correlation was obtained for $l=0.086$ and $f=1.2$. The correlation coefficient is then $r=0.69$.

For the *T-W* data the analysis showed that additional variation of other parameters was worthwhile. As a first approximation the following values were obtained:

$$\begin{aligned}\alpha &= 0.0053 \\ \beta &= 0.014 \\ p &= 1 \\ f &= 1 \\ l &= 0.05\end{aligned}$$

The main differences of the *D-F* and *T-W* data, compared to the Chickering optimization, are:

1. the relative weighting of luminance variations is considerably lower for tolerance data (parameter l);
2. tolerance data indicate clockwise rotation of chromaticity discrimination ellipses as compared with the MacAdam ellipses ($f > 0$).

Then the FMC-1 metric too should be abandoned in future work on tolerance data. The parameter values in the MacAdam modified 1965 Friele metric should be adapted to the data in question. In case this cannot be done it is advised to use the FMC-1 metric, however taking $l = 0.08$ and $f = 1.0$. Further development is urgently needed as even an adapted FMC metric will not turn out to be foolproof for the following reasons:

1. Fig. 5 shows that the FMC-1 metric with modified parameter values for l (0.086) and f (1.2) yields ΔE values for the 50% acceptance level which show a considerable and systematic variation. There is a tendency for lower ΔE values for grayish colours than for saturated colours. Fig. 5 is valid for the *D-F*-data.

The *T-W* data show the same tendency and the higher ΔE values are again

about twice as high as the lower ones. The effect is possibly caused by contrast.

2. The question whether the size of chromaticity ellipses for equal tolerance should depend on luminance (FMC-2, Simon-Goodwin) or not (FMC-1) was answered before in the negative way, as the data give no support for luminance dependency. McLaren (1969) discusses this point and states that a chromaticity difference which is readily perceived when sample and standard are white will be undetectable when they are black. This is true of course. It was only argued that this tendency is not distinct in the luminance range covered and certainly cannot be handled in the way as done in the S-G system and in FMC-2. For blackish colours the effect must be present. The change from Weber to de Vries-Rose discrimination is equivalent to a swelling of the chromaticity ellipses, however not as a mere multiplication.

It is hoped that further development of the fluctuation theory will handle this aspect in a correct way.

REFERENCES

- Billmeyer, F. W. (1969), The MacAdam color difference metrics. *Optical Spectra* 3 (6), 64-70
- Brown, W. R. J. (1951), The influence of luminance level on visual sensitivity to color differences. *J. Opt. Soc. Amer.* 41, 684-688
- Brown, W. R. J. (1957), Color discrimination of twelve observers. *J. Opt. Soc. Amer.* 47, 137-143
- Brown, W. R. J. and MacAdam, D. L. (1949), Visual sensitivities to combined chromaticity and luminance differences. *J. Opt. Soc. Amer.* 39, 808-834
- Chickering K. D. (1967), Optimization of the MacAdam modified 1965 Friele color difference formula. *J. Opt. Soc. Amer.* 57, 537-541
- Chickering, K. D. (1971), FMC color difference formulas: clarification concerning usage. *J. Opt. Soc. Amer.* 61, 118-122
- Davidson, H. R. and Friede, E. (1953), The size of acceptable color differences. *J. Opt. Soc. Amer.* 43, 581-589
- Friele, L. F. C. (1961), Analysis of the Brown and Brown-MacAdam colour discrimination data. *Farbe* 10, 193-224
- Friele L. F. C. (1965), Further analysis of color discrimination data. *J. Opt. Soc. Amer.* 55, 1314-1319 *Farbe* 14, 192-204
- Friele, L. F. C. (1966), Friele approximations for color metric coefficients. *J. Opt. Soc. Amer.* 56, 259-260
- Friele, L. F. C. (1969), Preliminary analysis of the Munsell colour system in terms of the Müller theory. *Proc. Int. Colour meeting "Color 69" Stockholm 1969*, vol. 1, 275-290
- Friele, L. F. C. (1971), Color difference and color tolerance evaluation. Problems and outlook. *J. of Materials* 6, 755-765
- Kuehni, R. (1970), The relationship between acceptability and calculated color differences on textiles. *Color Eng.* 8 (1), 47-53
- MacAdam, D. L. (1942), Visual sensitivities to color differences in daylight. *J. Opt. Soc. Amer.* 32, 247-274
- McLaren, K. (1969), Scaling factors in color difference formulas. *Color Eng.* 7 (6), 38-44
- Schultze, W. (1969), Umfassender Vergleich von sieben Farbabstandsformeln. *Proc. 1st AIC congress "Color 69", Stockholm*, 621-640. *Farbe* 18, 105-124
- Schultze, W. and Gall, L. (1969), Experimentelle Überprüfung mehrerer Farbabstandsfor-

meln bezüglich der Helligkeits- und Sättigungsdifferenzen bei gesättigten Farben. Farbe 18, 131-148

Thurner, K. and Walther, V (1969), Untersuchungen zur Korrelation von Farbabstandsbewertungen auf visuellem Wege und über Farbdifferenzformeln. Proc. 1st. AIC congress "Color 69" Stockholm, 671-687 Farbe 18, 191-206

Vos, J. J. and Walraven, P. L. (1972), An analytical description of the line element in the zone-fluctuation model of colour vision I, II. Vision Research, 12 in press

DISCUSSION

Wyszecki: Would it be possible to plot in your Figs 3 and 4 also the experimental points you used to derive the hypothetical lines?

Friele: The plots of the experimental points are available and will be added in the proceedings.

Jaeckel: Fkk-metrics correlates better with some data than FMC does, not only in overall correlations but for some individual colour centres as well. The relevant visual instrumental correlation coefficients, $x(-100)$, for the HATRA-435 group were:

Hue No. of Patterns	Blue 91	Red 91	Brown 94	Yellow 34	Green 125	All 435
Fkk	83	76	48	87	56	66
FMC	69	77	46	78	36	51

Friele: Although this may be true, it remains that Figs 5 and 6 of my paper do not favour Fkk and therefore principally Fkk should be abandoned. Furthermore the individual correlation is likely to alter when the parameter values are adapted to the data in question. For instance the lightness contribution is exaggerated in FMC-1 and would be reduced to $1/3$ or $1/4$.

Simon: How would you scale lightness for FMC-1?

Friele: This has to be done in a logarithmic scale at least down to $Y=10\%$. The Davidson-Friede data are not in conflict with a logarithmic scale. An analysis of the lightness contribution per colour centre for the Thurner-Walther data did not show systematical deviations from a logarithmic scale either.

28 GENERAL DISCUSSION II

Bartleson, Discussion leader

Bartleson

This has been an unusual meeting.

We have addressed ourselves to one particular subject from different directions: We have heard *modelmakers*, who were concerned with the development of viable and reasonable models to describe quantitatively an ideal color space. We have heard from the *experimentalists* who contributed new results for that purpose. We have heard from the *empiricists* who simply want to collect data and describe them in an useful way. Finally we have heard from the people on the firing line in *industry* and who must deal with these problems daily.

Quite a few problems have emerged from this meeting:

- Can we understand perceptual differences?
- Is there a difference between perception tolerances and acceptance tolerances? It seems that this depends on the application—but even this has been questioned.
- The utility of various color difference formulae is somewhat questionable—the various contradictory results being rather confusing.

So the rest of this afternoon may be devoted to: Where we ought to go next?

McLaren

It has been mentioned earlier, this conference, that, when people collect data, more often than not their own formula comes out on top. However, an unbiased approach is possible. When I first completed the Davidson-Friede analysis, the FMC formula came out best. However, as soon as I applied tests of significance any formula came out just as good or as bad as any others, as I showed in my Stockholm paper. Today, I was extremely impressed by the evidence in favour of the Adams/Nickerson type of formula. I was rather surprised that Prof. Simon could only quote that he had found that the MacAdam ellipse formulae—of the Simon-Goodwin type—were better for reds and purples. I get the impression that he has never done the comprehensive work of the type Mr. Jaekel has reported. In the U.K. we have provided evidence in favour of the Adams/Nickerson formulae. Could I ask whether there is any evidence available—I am sure it is not published—to support the MacAdam type formula?

Simon

In the earlier days—you are absolutely right—there was not this comprehensive work, we had no means to do it. I must defer certainly to Mr. Jaeckels work. The body of data—small as it was—is completely unavailable to me. I don't know where it is. It was years ago, I only vaguely remember what we did.

Ishak

There are two main weak points in our approach.

1. On the input side: the standard observer is questioned.
 2. On the output side: there is great difficulty in expressing the perception.
- So what we need is more work on the standard observer and more work on psychological scaling.

Jaekel

I am sure there are some things wrong with my data—just as there are some things wrong with everybody's data. Perfection simply takes a long time. As far as representative assessors are concerned, we had 5 people who took part in all the three groups: 18 people who took part in two of the three groups and 29 people who took part in one of the three groups.

I know, it appears asymmetrical, but it totals up to 52 people, which is a reasonable number of experienced industrial assessors.

Then a second point. When I talked about the probability of wrong decisions, I based my calculations on the assumption that you use the same tolerance for all hues. Almost all the papers here—and our own more detailed analyses—seem to indicate that you may have to use different tolerances for different hues. Obviously the superiority of the instrumental approach to the visual one regarding wrong decisions is greatly to be enhanced, then, if it is refined, when required, by tailoring tolerances to hues.

Clarke

Coming back to Dr. Ishak's first point: The standard observer, right or wrong, does not matter as long as you are concerned with small differences with no metamerism. He only matters for metamerism and colour matching.

Brockes

I might report on another method to come to a useful color difference formula, which is used by Rolf Kuehni in USA and will be published by him. Instead of taking an already known color difference formula and merely checking it by correlation with visual data, he tries at first to draw ellipses which represent the visual acceptance data around each standard as good as possible. There is much argument in which way this can be done, and another approach as that of Kuehni may easily be better justified, but this is a detail which should

be of no concern here. Once we have such ellipses, the second step is to optimize a color difference formula to this set of ellipses or better even ellipsoids. This may be done in a similar way as Dr. MacAdam has set up the Xi-Eta-formula, but also formulas of other structure may be used as e.g. the Friele-type formulas, if they only have enough parameters for the optimizing process. In this way an improved formula can be found. I believe that it is really important not only to check old formulas which base on unproven models of vision, but also to construct new formulas which base directly on visual data for acceptability. In this way—by setting up first experimental acceptability ellipsoids—it is also possible to compare directly shape and direction of these ellipsoids with the threshold perceptibility ellipsoids.

Bartleson

Dr. Wysecki indicates that he has a copy of Kuehni's paper (Kuehni, 1971).

Wysecki

Just before I left for Europe, I had a telephone conversation with Mr. Kuehni and he told me about his new work on the problem of acceptability and perceptibility. He has earlier pointed out that there is a distinct possibility that assessors in the textile industry are biased. They introduce criteria other than those of strict perceptibility in assessing the differences they see. Specifically, they introduce acceptability criteria. Hue differences were not tolerated as much as other differences.

Now he has produced loci of acceptability from the Davidson/Friede, Robertson, his own, and "Metropolitan Section" data. His method is complex, but it is interesting to see that the loci all turn out to be ellipses. In the chromaticity diagram they distribute not always in a regular fashion, but with a sudden break in some areas. Kuehni's conclusions are not always those I would have drawn. The observational data are wonderful clouds; to make statistical inferences from them is most difficult, to say the least.

Most ellipses point to a part of the spectrum locus of higher wavelength than the MacAdam ellipses. The difference in orientation being of the order of 20° – 30° . There is a tendency for colors of greater saturation to have increasingly larger acceptability ellipses. To my opinion Kuehni has made an important contribution to the problem of acceptability versus perceptibility, but the interpretation of his results seems to be in difficulty and not consistent.

Brockes

May I comment to this. I discussed this also by letter with Kuehni and I share your opinion completely. I think we should not discuss here so much his particular ellipses, but rather his approach. I should like to have your opinion on that.

Bartleson

Time is running out. May be we can get together tonight to continue this interesting general discussion?

REFERENCES

Kuehni, R. (1971), Acceptability contours and small color difference. *J. Color and Appearance 1*, 30-35

29 GENERAL DISCUSSION III

Walraven, Discussion leader

Walraven

We have got together another time to discuss some points not adequately answered so far. I suggest Dr. Wyszecki to take up the thread of this afternoon.

Wyszecki

I should like to have your opinion on the following three points which would be very helpful for future discussions at Barcelona. The CIE Colorimetry Committee has to come to grips with the massive amount of new and old data we have before us.

1. I would like to know whether there is a conceptual difference and a difference in fact between acceptability and perceptibility studies. Would it be possible to decide on the kind of studies we should make to resolve this problem?
2. Can loci of constant perceptibility—or acceptability be assumed to be of ellipsoidal shape having certain sizes, orientations and eccentricities?
3. In all these studies, have we taken sufficient account of the uncertainties in the physical calibrations of the samples and stimuli and the enormous spread in the observational data, and how much can we rely on statistical inferences? How significant are significance tests?

Nimeroff

I have missed in our discussions the term vector. We talk about color-differences as if they were only scalar, but they have both size and direction! We have here emphasized only the magnitude aspects and neglected the important directional aspects.

Vos

I don't understand your problem. Did not we talk about ellipses, with size and direction —so did not we discuss vectors, be it under another name?

MacAdam

A vector that depends on both location and direction—that is what we call a tensor. The ellipses, or in general ellipsoids, portray the perceptibility of color

difference (or, operationally, the variance of color matching) as a tensor. But, I don't think we need get involved with such semantic niceties or with the mathematics of vectors or tensors. The minimum requirement that is not now satisfied and does not seem likely to be satisfied in the near future, is a generally accepted formula or method for calculating a single number to specify any color difference. If we had that, its usefulness would not be destroyed if someone called it a scalar. If we had it, we could then consider how to resolve it, i.e., break it down into its components—hue, saturation and brightness, or any other triad you prefer. But, first of all, we need to agree on a single formula for color difference that we will all use, and not fly off in all directions deriving, or using, others.

Wyszecki

I don't feel a need for elementary mathematics. Cannot we go back to perceptibility vs. acceptability? There is a difficulty in what we mean by acceptability and by perceptibility. Mr. Jaeckel, for instance, seems to have a definition that deviates from others.

Jaeckel

I think I have made quite clear what is my definition.

I state in my paper that assessors were asked to answer with "yes" or "no" the question "Would you accept this pattern as a commercial match to this standard pattern?" Thus I define acceptability as the percentage of "yes" comes in the total answers obtained for a pattern, and I think this does not deviate at all from other published definitions of acceptability.

When there is no perceptible difference, there is no difference between acceptability and perceptibility: acceptability is 100%. When there is a large perceptible difference, there is little or no difference between acceptability and perceptibility: acceptability approaches or equals 0%. Small perceptible differences may or may not be acceptable, depending on whether the difference is mainly in reflectance or mainly in chromaticity for instance, or on whether the assessor who perceives the difference is lenient or severe. Always, however, acceptability is related to perceptibility (often closely), which should be measured by colour-difference formulae: it is possible to overstress the difficulty of differences between acceptability and perceptibility.

Walraven

So that this can not be a problem anymore. I think we are on the wrong track if we continue this way.

I doubt in particular whether we will come much further by wrangling about acceptability and perceptibility. Should we not leave that dispute for the moment and restrict ourselves to perceptibility?

Unknown

I agree completely. Let us realize: perceptibility is something which has to do with the eye. Acceptability involves a factory—customer relation with all intricatenesses involved. It is a swamp in which we are going to sink away.

Clarke

I should like to come back here to the Walraven/Vos model, which I admire for its elegance, but which, I think, is fundamentally wrong—and that for the following reasons.

Firstly we have the theme of the eye operating as an ideal physical detector system, a theme recurrent in the works of Rose, de Vries, van der Velden and Bouman. As the model is presented here to explain colour discrimination, let us first consider chromaticity discrimination. For a 2° bipartite field the eye can just see a difference of about 0.001 in the white region of the (x, y) chart under best illumination conditions, or around 0.001 of the size of the gamut of real colour stimuli. If the eye's receptors had spectral responsivities like $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ then this would imply that the relative uncertainties $\Delta X/X$, $\Delta Y/Y$, $\Delta Z/Z$ in the X , Y , Z responses would be of order 0.001 also. We know that the spectral responsivities $\bar{p}(\lambda)$, $\bar{d}(\lambda)$, $\bar{i}(\lambda)$ of the three cone systems are not all that dissimilar to $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$, and hence in the colour space (P, D, T) the colour discrimination will be of the same order giving $\Delta P/P$ etc. the value of around 0.001. This is valid for the transformation from (X, Y, Z) to any of the more plausible response systems that have been proposed, including the (R, G, B) system of Vos and Walraven. This implies that the number of useful quantum absorptions in each of the three cone systems must be of order 10^6 since $\Delta N/N = N^{-1/2}$.

Unfortunately, photometric calculation shows that the number of useful quantum absorptions is far higher than this under the viewing conditions for best colour discrimination. Assuming 3 mm pupils, Le Grand's data on the schematic eye, a sunlit illumination of $1.0 \times 10^5 \text{ lm/m}^2$ falling on a light grey sample of 65% luminance factor, and the commonly accepted value of 10% quantum efficiency for the situation when the receptors involved are densely packed on the retina, then we calculate that there must be around 3×10^{10} quanta/sec usefully absorbed for each half of a bipartite 2° field. A viewing time of at least a second is needed for best discrimination, so at least 3×10^{10} useful events must take place. This means that about 3×10^4 times more quantum absorptions are needed by the eye than by an ideal physical detector! The situation is even worse if we consider the remaining aspect of colour discrimination, that is, luminance discrimination, for the eye can only see about a 1% difference under best conditions (which cover most conditions met with in photometry and colorimetry). This implies that only 10^4 or so useful events are needed, so that the discrepancy is a factor of up to 10^6 for high levels of

illumination. For a 10° field, the number of events involved is of order 10^{12} , and as discrimination is almost 3 times better than with a 2° field it is clear that the discrepancy between eye and ideal detector is of the same order for these conditions also.

Because of the gross shortcomings of the ideal detector theme at high levels, Vos and Walraven added the concept of a "photon counter" (see p. 75 of their paper) in an attempt to get better agreement with visual performance. Unfortunately, this does not account for even the basic facts. If it were a *true* photon counter it would affect the photon initiated events, implying that a "dead time" of order 10^{-11} sec was involved (since optimum discrimination obtains at 10^{10} to 10^{11} quantum events per sec).

These times are not credible for any known physiological mechanism that could be supposed to participate: the shortest plausible dead time would be of order milliseconds.

This model and its predecessors fails to account for the most significant fact of visual discrimination, the Weber-Fechner Law. This is the most general law, which is true under almost all conditions that are important, i.e. those met with in practical photometry and colorimetry work: The failures of the model in various respects (which we have not time to discuss this evening) can be traced mainly to the fact that the model predicts square-root law behaviour (except at the highest level of the photon counter saturation) but that under most conditions the threshold versus luminance curve shows either a very limited range of square-root behaviour or often none at all. Extensive experiments by Stiles and many others before and after show this to be so. In the model Weber-Fechner Law behaviour must represent a non-ideal situation due to overloading. Unfortunately, the theory of a counter with some assumed "dead time" shows that the overloading would only be significant for about a log unit below the level of saturation whereas the eye has many log units of Weber-Fechner behaviour extending below its saturation region, the region where discrimination is reduced with further increase of luminance. While I could raise the questions of other weaknesses in the model, I feel the problems I have mentioned are weighty enough to raise serious doubts as to the usefulness of the model, and we only have a limited time for discussion tonight.

Boynton

Before you answer Clarke, I should like to add a comment about a further problem with the model. In California, I worked with David Whitten in Kenneth Brown's laboratory, on the late receptor potential. (Boynton and Whitten, 1970). We found a nonlinear relation between stimulus intensity and response magnitude, which is one of two adaptational mechanisms that appear to exist in the photo-receptors, and which do not depend upon feedback.

Vos

Clarke's arguments consist of facts—which are not all correct—and an interpretation of our model—which is too much simplified.

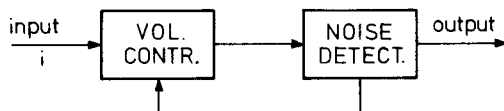
First the facts. You compare 10^6 quanta, needed if the eye were an ideal detector, with the experimental value of 3×10^{10} quanta/sec, 2° field. That is no fair comparison, though, as it seems dubious whether 1 sec and 2° are not much too large to act as sampling units. I admit that we did not explicitly mention the term "sampling unit" in this paper, so that misunderstanding might arise. From the Malkin-discussion we have learned another time that it is the border region and not the whole field, that is crucial. Further, your level of 1% for luminance discrimination seems too high. Under suitable conditions, values of 0.3% have been found. Finally, the expression $\Delta N/N = \sqrt{N}$, should of course have a constant k in it ($\Delta N/N = k\sqrt{N}$), k being dependent on the degree of certainty with which a luminance or colour difference statement is made. To my opinion $k = 2$ or 3 might be a better guess in this respect. That does not take away your main point, but these effects may greatly reduce the discrepancy. Therefore we agree that our model is not a final detailed realistic description of what happens in the visual system. We have rather investigated how an ideal contrast processor would behave and have then gradually introduced "imperfections" like limited channel capacity in a most simple version. And then we looked how far we could predict the actual overall behaviour. As to that we feel to have managed quite well.

The most simple imperfection introduced was the dead time which sets a limit to the strength of the output signal. First: we know that this is not a realistic model in that these effects probably occur before spikes occur at all. But it has the advantage that we can quantitatively describe with it a process of the sort which should happen anyway. Refinements are possible—and are suggested. Let us take the square-root scaling as suggested by Bouman and Ampt (1965) and by Barlow (1965). With a dead time of 10^{-3} sec we expect overloading above 10^3 events/sec. input; with square-root scaling that would occur above 10^6 events/sec. input.

Effectively this comes down to a dead time of 10^{-6} sec as Trabka (1969) has shown. Can not this be in the order of your corrected value?

Since this down scaling would only affect the region of overloading, but not the accuracy (since no information is lost) our model is not essentially affected in this way apart from in the value of the dead time.

In fact in our present thinking the principle of down scaling to constant noise level is very attractive even at higher levels.



If the visual system would apply a volume control by feedback so that the output has always constant noise level (like we are to apply volume control on a noisy scope signal until the grass is just near the limit of detectability) the output would go up with the $\sqrt{i}$ for low levels, and with $\log i$ for high levels, where overloading occurs. This would bring us, by the way, quite near to the multiplicative definition of V_d , as by Stiles.

Clarke

I am sorry, but your explanation goes just the wrong direction, because it would delay overloading and then Weber's law, where we know Weber's law to hold at all photopic levels.

If you assume square root transformation before the "dead time" mechanism, you would need a dead time of not much more than a microsecond. Even this time is not physiologically acceptable.

Moreover: Your additional information makes almost a new model!

Boynton

I have also thought often in terms of feedback "volume control," but have felt compelled to abandon the idea in view of our findings with the late receptor potential.

Walraven

I will underline Vos' statement that it is not fair to compare the discriminative possibilities in terms of the total number of quanta in the test field with the actual measured values. The eye uses the incoming information for visual acuity which corresponds with smaller angles than a 2° field. The sampling units to be used in the fluctuation concept are much smaller than the usual field sizes in photometry and colorimetry. If one takes the discriminative power in space and time also into account, one sees that the de Vries-Rose law is the law, and the the Weber-Fechner law should be considered as a secondary law as the van Meeteren/Vos paper (1972) shows.

Further, in colour discrimination the de Vries-Rose law is not limited to a marginal luminance domain. In fact I have shown that it holds over 3 to 4 log units when at equal luminance a test spot of the size of a few arc minutes up to one degree is compared with a large surround (Walraven, 1967).

Clarke

Foveally measured with good fixation?

Walraven

Yes, indeed.

Vos

May I give the discussion a bit another direction, since I guess not all participants can directly evaluate the numerical evidences pro and con. I should like to give the discussion sort of philosophical twist.

When we, of the organizing committee, planned this meeting, we tried to get the symposium theme gradually moving from more basic on the first day to more practical on the last day. We had hoped that the more theoretical papers might give a basis on which to understand the practical results. I must confess to be a bit disappointed that there has been so little reference to the theoretical papers: that our model, for instance, seems to be sort of silly study in retrospect. Now, I would not object to this situation if the practical people would have shown to be clearly able to manage without fundamental knowledge. However, my main impression this last day in particular is that they terribly need us. Dr. Wyszecki has politely talked about a mass of data, I would prefer to speak of a mess. If that is the right description, do not we need more theoretical thinking? May be our model is not the answer, but it is a proposal, a challenge to say, and the most detailedly predictive one there is, for that matter. And is not that the main role of a model? May I give an example from history. Einthoven explained in 1885 the colorstereoscopic effect in terms of the lateral position of the fovea, though he knew that he was only half correct at best, since 50 % (!) of his observers had the effect the wrong way round. Brücke had described the same mechanism 17 years before, but abandoned the explanation as it did not fit all the experimental data. Brücke's work became forgotten, whereas Einthoven's work became the stepping stone for further understanding.

Stiles

I am not so sure that Dr. Clarke and Dr. Boynton's objections against the Walraven/Vos model are as strong as they suggest them to be. I have followed the gradual development of the de Vries/van der Velden idea's via Bouman/Walraven/Vos and it seems to me a stimulating attack on the problem. It may need modification — perhaps radical modification — but it exemplifies the kind of inductive approach which I believe fruitful.

Walraven

May I suggest we try to turn to Dr. Wyszecki's last question: What experiments are we going to do?

Stiles

Apparently there is a gap between the MacAdam ellipses, determined on "aperture-colours" and the industrial tolerances on tiles and textiles. Could not we bridge the gap by repeating the MacAdam/Brown type of experiments

—may be for only a few checkpoints—for both aperture and texture colours? I think that would provide a more firm basis for future discussions than we now have.

MacAdam

I wholeheartedly agree, and it underlines what I have time and again said during this conference: do the experiments. I will not be able to do such work anymore, but let others come into it. As to the difference between aperture and surface colors, I must say that our experiments were not that abstract. Both the PGN and the Brown 12-observer studies were made with bright, daylight-quality surrounds, sharply contiguous with the test fields, which produced modes of appearance such that you could not have told the difference of the test fields from illuminated sheets of paper.

Vos

I have another question—though it may be that I am asking silly questions; I am only shortly working in colour. We have talked about acceptability and perceptibility. Now acceptability is a question of judgment—a typical cortical function to my opinion. Perceptibility may be a question already determined at the retinal level—just for economy's sake. If that is true, it is not necessary to assume much relation between the two—apart from the fact that one is a minimum condition for the other. A check on this thesis might come from color discrimination with the comparison fields in two eyes. Has that been done?

Hunt

Yes, we did. The accuracy of the hetero-ocular experiments is way down: compared with the ordinary experiment. A factor of 10, roughly.

Boynton

Let's put the problem in more general terms. The perceptibility of color difference depends upon many things. In the first place, if we lived in a world illuminated by monochromatic light, or which contained only gray surfaces with "flat" spectral reflectance curves, there would be no physical basis for color perception. Given the physical basis which actually exists (a continuous and balanced spectrum of illumination, and surfaces with widely varying spectral reflectances), there could be no color discrimination without photo-sensitive mechanisms capable somehow of discriminating among various spectral distributions. We know how this works now, being based upon three classes of cones, each responding in accord with Rushton's Principle of Univariance, but differentially tuned with respect to spectral sensitivity. But this initial "color separation" is not sufficient, since if the optic nerves were severed there

would be no vision: the initial trichromatic information must be further processed, delivered to the visual cortex, and so on.

Stiles

One must also remember the final term in the relation. The cortex is not the end point. There must follow some overt response in the physical world. The man has to turn the knob of the monochromator, for instance.

Vos

That is an important point. It may quite well be that the controversy about the Parra experiments is determined by this human factor problem. And then, we can leave the infertile acceptability vs. perceptibility controversy since acceptability so clearly involves the customer-factory relation.

In the beginning of this discussion I plead for more fundamental work to get to understand colour discrimination. May be I should complement that statement now by asking more fundamental psychological research on tolerant-ness and related problems if one is interested in industrial acceptability. To talk about acceptability if we don't know what that is seems wasted time to me.

Walraven

Dr. Wyszecki has not exactly got his answer, but I think the central issues of this symposium have come quite a bit more clearly outlined at this final round of discussions, and if we can make a conclusion regarding the work of the CIE it seems that that organization has to limit its work to perceptibility.

Let us go to Barcelona.

REFERENCES

- Barlow, H. B. (1965), Optic nerve impulses and Weber's law. Cold Spring Harbor symposia on quantitative biology, vol. 30: Sensory receptors, pp. 539-546
- Bouman, M. A., and Ampt, C. G. F. (1965), Fluctuation theory in vision and its mechanistic model. Proc. symposium "Performance of the eye at low luminances", Delft, pp. 57-69
- Boynton, R. M., and Whitten, D. N. (1970), Visual adaptation in monkey cones: Recordings of late receptor potentials. *Science* 170, 1423-1426
- Meeteren, A. van, and Vos, J. J. (1972), Resolution and contrast sensitivity at low luminances. *Vision Res.*, 12, in press
- Trabka, E. A. (1969), Effect of scaling optic-nerve impulses on increment thresholds. *J. Opt. Soc. Amer.* 59, 345-349
- Walraven, P. L. (1967), A uniform chromaticity diagram based upon a square root transformation of the colour space. Proc. 16th meeting CIE, Washington, Vol. A: 106-111

30 RECENT DEVELOPMENTS ON COLOR-DIFFERENCE EVALUATIONS*

GUNTER WYSZECKI

*National Research Council
Ottawa, Ontario, Canada K1A 0S1*

In this paper, the major developments on color-difference evaluations will be reviewed, beginning with the year 1964 when the Commission Internationale de L'Eclairage (CIE) recommended a color-difference formula for provisional use. The review is supplemented by a bibliography of over 200 references covering color-difference evaluations and related subjects from the period 1964 to 1971.

I. INTRODUCTION

The first objective of a colorimetric measurement is to ascertain a set of quantities that suitably describe either the property of light or the visual perception by which an observer may distinguish between two fields of view of the same size, shape and structure. Differences between the two fields of view may be caused by differences in the spectral composition of the light observed.

Suitable quantities that describe the colorimetric property of light are the tristimulus values, suitable quantities that describe the visual perception of an observer viewing a patch of light are hue, saturation, and brightness.

The determination of the tristimulus values of a patch of light (color stimulus) rests on the ability of the observer to match uniquely the color of the given stimulus with the color of another stimulus produced by an additive mixture of three primary stimuli. The laws of additive mixture of color stimuli are well established. They have been utilized fully in the development of the colorimetric system recommended by the Commission Internationale de l'Eclairage (CIE).

The determination of the quantities hue, saturation, and brightness, per-

* Added to these Proceedings on invitation.

ceived by an observer viewing a given color stimulus, rests on the ability of the observer to isolate these quantities subjectively. Hue, saturation, and brightness are initially undefined in any quantitative or numerical sense, and it is necessary to construct scales from purely subjective judgments made by the observer in extensive experiments.

The subjective judgment that characterizes the observer's perception with regard to a given color stimulus is influenced strongly by the conditions of observation. First there are the physical characteristics of the test stimulus, such as its size, shape, luminance and relative spectral power distribution. Then there are the other color stimuli which are in the field of view surrounding the test stimulus in some given configuration. Their sizes, shapes, luminances and relative spectral power distributions will affect the observer's judgment on what color he perceives for the test stimulus. Movement of the stimuli or the observer's eyes and exposure times of the stimuli are other factors to be considered. Indeed the problem is a multi-variate problem of enormous complexity and to a certain extent it is still only barely understood.

Nevertheless, a colorimetric method that aims at a quantitative measure of the visual perception of an observer, exposed to a display of different color stimuli, is perhaps the ultimate and most desirable from the theoretical as well as practical point of view. However, such a method does not exist as yet.

The colorimetric system of the CIE which yields tristimulus values for a given color stimulus was never intended to provide—and indeed does not provide—information on what color an average observer would perceive for the stimulus. All that the CIE colorimetric system offers, within a certain range of conditions of observation, is a prediction of whether two given color stimuli of different spectral power distributions will be perceived to have the same color. There is an overwhelming evidence that the predictions made by means of the CIE colorimetric system agree remarkably well with visual assessments made by observers with normal color vision, and many industries have adopted the system to assist them in their color-production-control tasks.

The second objective of a colorimetric measurement is to ascertain a quantity that suitably describes the color difference an observer may perceive between two given stimuli. The determination of such a quantity rests on the ability of the observer to judge the relative magnitude of two color differences he may perceive when viewing two pairs of stimuli. The observer's judgment varies greatly with the conditions of observation and the kind of stimuli presented to him. The problem is closely associated with that of determining quantities that suitably describe the observer's perception with regard to a single stimulus viewed in a display of different color stimuli. Sizes, shapes, luminances and relative spectral power distributions of the test stimuli and those surrounding them are again important parameters affecting the observer's judgment.

The precision of an observer's color-difference judgments, assessed in any reasonable way, invariably compares poorly with that of strict color-matching. The difficulty of making such judgments is reduced if the stimuli concerned give rise to color perceptions closely similar in one or two of the three basic attributes of color perception: hue, saturation, and brightness. Thus judgments of the relative magnitude of brightness differences are readily made and with good precision, if the stimuli viewed have at least approximately the same hue and saturation, that is the same chromaticness. Similarly, judgments of differences in chromaticness, if the stimuli are perceived to have the same brightness, are not generally found to be too difficult.

As the brightness and chromaticness differences to be judged become larger, the observer finds it increasingly difficult to assess their magnitude. There is also a very limited practical value in judging extremely large color differences. Of principal interest are the medium-large and small differences and, in the limit, those differences which the observer perceives to be *just noticeable* or to be *on the threshold*. A just noticeable or a threshold difference is considered the natural unit for measuring larger differences. However, the visual phenomena and concepts associated with threshold judgments differ in various ways from those involved in the assessment of small but clearly perceived color differences. It is quite conceivable, that in the latter case the action of the cortex, as part of the visual mechanism, becomes a predominant factor in the observer's judgment and contributes increasingly to the variability of the results. In extreme cases of judging very large color differences, the results become meaningless.

Different approaches have been taken to arrive at a method of predicting the color-difference judgment of an observer viewing two given stimuli under given conditions of observation. However, all approaches are characterized by the common concept that perceived colors can be represented by points in a three-dimensional space. The problem then is to measure distances in this space which correspond to perceived differences between colors. This introduces the mathematical notions of the *line element* and *color difference formula*.

Line elements and color-difference formulae provide methods for assessing color differences, but there are significant differences in the way they have been determined and in the way they are used. The color-difference formulae, in nearly every case, assume that color-perception space is Euclidean, or approximately so, but a line element can be used even when this assumption is not valid. The basic observations used in setting up line elements are threshold measurements and standard deviations in color matching. The observations underlying the color-difference formulae are related mainly to small but not actual threshold differences and the formulae are proposed for application to both larger and smaller differences.

All line elements put forward up to now have been assumed to have the

Riemannian form defined by a definite positive quadratic equation for the *just noticeable* color difference, ds , or a constant fraction of it. The outstanding problem in developing a line element of this kind is determining the coefficients of this quadratic equation, the so-called *metric coefficients*.

Two differing approaches have been followed in determining the metric coefficients. One is based on theoretical considerations regarding the functioning of the visual mechanism coupled with certain experimental threshold data. This approach may be described as the *inductive method*.

The other approach may be called the *empirical method*, in which the metric coefficients are derived by an empirical analysis of large blocks of measurements of threshold differences, or of the closely related standard deviations of color matching, obtained for colors covering an extended domain of color space.

Several important color-difference formulae are based on *color-order systems* which have been constructed on principles of color perception. Color chips, displayed on a white or gray background and illuminated by daylight, have been selected to represent scales of constant hue, saturation, and lightness. Each scale is spaced uniformly in accordance with the perceptions of an observer with normal color vision. The Munsell system is an important example of such a color-order system.

Each color perceived in the ordered collection of color chips can be represented by a point in a three-dimensional space where hue, saturation and lightness form a cylindrical coordinate system. In a plane of constant lightness, loci of constant hue are radial lines starting at the point representing gray; equal angles between these radial lines correspond to equal steps of hue. Loci of constant saturation are circles centered at the gray and are evenly spaced corresponding to uniform steps in saturation. A plane of constant lightness exhibits points representing a chromaticness scale in a polar-coordinate system. The complete three-dimensional space is constructed by combining the chromaticness scales of different lightness levels such that their respective central grays are uniformly spaced between black and white.

The particular color chips, that realize the graduations of the perceptual scales of the color-order system, can be measured in terms of CIE tristimulus values with respect to the CIE standard colorimetric observer and an illuminant representing daylight. Analytical expressions can then be sought that transform these tristimulus values to three new variables which, when used as rectangular coordinates, define the desired color-perception space. This is the space in which a given perceived difference between any two color chips corresponds to the same distance separation of their representative points. With the Munsell system used as a base, several attempts at simple but usually nonlinear analytical transformations have been made with some success.

In a somewhat different approach to perceptually uniform scaling, the CIE

(x, y) -chromaticity diagram is subjected to a projective transformation that would yield a new diagram in which equal distances correspond to perceptually equal color differences. In such a diagram only colors of constant luminance are considered and the term *uniform-chromaticity-scale diagram* (= UCS diagram) has been used for it. UCS diagrams prove to provide only crude predictions of perceived color differences, but the simplicity of transformation from (x, y) -chromaticity coordinates is considered a valuable feature in some practical applications. The limitation of a UCS diagram to colors of constant luminance can be overcome by combining it with a lightness scale, such as that employed in the Munsell system.

Many line elements and color-difference formulae have been developed and several are being applied to color-production-control problems in industry. However, none has found universal acceptance among different industrial laboratories.

Introductory accounts on operational details of the measurement of tristimulus values, on color-order systems, such as the Munsell system and others, on color-difference formulae and line elements can be found in recent books, for example, Judd and Wyszecki (1963), Graham et al (1965), Billmeyer and Saltzman (1966), Wyszecki and Stiles (1967), Hunt (1967), Sheppard (1968), Wright (1969), MacAdam (1970), and Cornsweet (1970).

2. THE CIE RECOMMENDATIONS

A significant event in the search for improved methods of evaluating color differences was the decision of the CIE colorimetry committee, in the late 1950's, to put the problem on their working program. This meant that experienced workers in colorimetry from different countries would jointly consider the problem and work toward its solution. This is the committee which had already produced the basic CIE recommendations on colorimetry concerning the standard colorimetric observer and coordinate system, standard illuminants, standard of reflectance, and standard illuminating and viewing conditions.

It was soon recognized that the task was a most difficult one, and the lack of new experimental data on color discrimination was the most important factor that prevented the committee from finding a satisfactory solution. However, the urgency of the problem was also recognized, and in 1960 it was decided to recommend provisionally the use of the UCS diagram originally suggested by MacAdam (1937). This particular diagram was chosen from among several similar diagrams because of the simplicity of the transformation that produces it from the CIE (x, y) -chromaticity diagram.

The provisionally recommended UCS diagram is known officially as the CIE 1960 UCS diagram. Its use is recommended whenever a chromaticity diagram is desired that yields color spacing perceptually more nearly uniform

than that of the CIE (x, y)-chromaticity diagram. The color spacing afforded by the CIE 1960 UCS diagram applies to the observation of color stimuli having negligibly different luminances. The stimuli, represented by colored objects, are of the same size and shape, and are viewed in white to middle-gray surroundings by an observer photopically adapted to a field of chromaticity not too different from that of average daylight.

In 1964 the CIE provisionally recommended an extension of the CIE 1960 UCS diagram to three dimensions, based on a proposal by Wyszecki (1963). The recommended rectangular coordinates U^* , V^* , W^* are nonlinearly related to the CIE tristimulus values X , Y , Z . The distance between two given points (U_1^* , V_1^* , W_1^* and U_2^* , V_2^* , W_2^*) in the CIE 1964 (U^* , V^* , W^*)-space defines a measure ΔE_{CIE} of the perceptual size of the difference between the two colors represented by the two points. The distance ΔE_{CIE} is formed simply by calculating the square root of the sum of the squares of the differences between the corresponding U^* , V^* , W^* coordinates of the two points.

The CIE recommendations of 1960 and 1964 were put forward as an attempt at unifying the diverse practice of evaluating color differences in industry and thus assisting many of those concerned with the problem of setting and describing color tolerances. At the time these recommendations were made, perhaps a dozen or so different schemes of predicting color differences were actually being used in different colorimetric laboratories. Such diversity of practice is not desirable. Color differences derived from one color-difference formula are difficult, if not impossible, to interpret in terms of another formula.

The CIE had only very limited success in its attempt at unifying the diverse practice of evaluating color differences. It was pointed out by various workers from industry that the provisional recommendations made by the CIE did not bring about the desired improvements in their particular work. The continuation of the search for improved methods of color-difference evaluation was urged and the enlargement of the working program of the CIE colorimetry committee was suggested.

In 1967 the CIE colorimetry committee recommended a detailed program for studying color-difference evaluations (Wyszecki 1968). This program includes the call for new visual experiments on assessing color differences. Some guidelines are given with regard to the observing conditions to be used in the experiments so as to conform with common industrial practice of visual color matching. The experimental data should then be used to check not only the color-difference formula provisionally recommended by the CIE in 1964, but also three other formulae under consideration as possible improvements of the CIE 1964 formula. The other three formulae are the Cube-Root formula, ΔE_{CR} , the Godlove-Munsell renotation formula, ΔE_{GM} , and the Friele-MacAdam-Chickering formula, ΔE_{FMC-2} .

3. REVIEW OF PROGRESS SINCE 1964

The CIE recommendations concerning perceptually uniform color spacing and the evaluation of color differences have increased the interest in the problem, particularly with respect to industrial applications. Although substantial progress toward a satisfactory solution of the problem may not have been made as yet, there is little doubt that a considerably deeper understanding of the problem exists now than ever before. All this is evidenced by the many publications bearing on the subject since 1964.

A bibliography is appended to this review. Authors have been arranged in alphabetical order. As a rule, only those publications are listed which fall in the period of 1964 to 1971. However, some references from the period prior to 1964 are added whenever this became desirable for the purpose of discussions included in this review. No claim of a complete bibliography is made, but it is hoped that no major contributions were missed.

In reviewing the published work of the various authors, the name(s) of the author(s) is (are) given with the year of publication. This information may then be used to find the complete reference in the bibliography.

The emphasis of this review will be on all those contributions that are specifically concerned with color-difference evaluations, such as line elements, color-difference formulae, tests conducted on them, and their application to industrial color-control problems. This part of the review will be under the main heading of *Color Metrics*. Other contributions, which have some bearing on color metrics and may be useful to considerations of future developments of color metrics, are briefly reviewed under the main heading of *Studies Related to Color Metrics*.

4. COLOR METRICS

4.1 *Line Elements*

The contributions by Friele (1965, 1966 a and b, 1967), MacAdam (1964 a and b, 1965 a and b, 1966) and Chickering (1967) are further developments of Friele's original proposal of a line element (Friele 1961).

Originally, Friele (1961) transformed the experimentally determined color-matching ellipsoids of Brown (1957) and Brown-MacAdam (1949) into a coordinate system based on the fundamental primaries of the visual mechanism. Friele found that the color-matching ellipsoids, which define the standard deviations in visual color matching with a tristimulus colorimeter, can be described adequately by assuming a Weber-Fechner type of differential sensitivity in the visual processes. The visual processes involved are assumed to comprise one summation process (lightness) and two antagonistic chromatic

processes (red-green and yellow-blue). When a small signal from one receptor mechanism is opposed to a larger signal from another receptor mechanism, the effective threshold of the first mechanism is limited by the "noise" of the second. Thresholds of the different mechanisms cannot be added as though they were independent. The assumptions made with regard to the functioning of the visual mechanism are closely related to the three-stage Müller theory.

In subsequent work on his line element, Friele (1965, 1966, 1967) revised the parameters of the line element so as to take into account also other experimental data on visual color discrimination. The other data included the color-matching ellipses observed by P. G. Nutting Jr., and reported by MacAdam (1942), as well as commercially acceptable color differences reported by Davidson and Friele (1953). In a series of papers, MacAdam (1964 *a* and *b*, 1965 *a* and *b*, 1966) pursued the optimization of the parameters of Friele's line element so as to obtain a best fit to the color-matching ellipsoids of Brown (1957) and Brown and MacAdam (1949), and the PGN color-matching ellipses (MacAdam 1942). As an additional result of these optimizations, the singularities in the g_{ik} -coefficients of Friele's version of the line element were removed and the g_{ik} -coefficients plotted as smooth curves in the (x, y) -chromaticity diagram.

A further optimization of the parameters of Friele's line element was made by Chickering (1967) along the lines suggested in MacAdam's work. Only the PGN color-matching ellipses were used in the calculations of optimizations because MacAdam's work, particularly MacAdam (1965 *b*), had indicated that the PGN ellipses were representative of young normal observers, and that these ellipses could be fitted by the line element better than other, similar data. The result of Chickering's optimization is now often referred to as the Friele-MacAdam-Chickering formula No. 1, abbreviated FMC-1.

In a subsequent development at the Washington meeting of the CIE colorimetry committee in 1967, already indicated in the introduction above, the FMC-1 formula was considered, among several others, as a possible improvement of the provisionally recommended CIE 1964 formula. However, the FMC-1 formula applies only to calculating differences between colors of the same luminance. An extended version of the formula was suggested by Chickering at that meeting, which included two additional parameters K_1 and K_2 , both functions of the luminance factor Y . This version, now known as FMC-2, was recommended by the committee for further study, together with three other formulae including the CIE 1964 formula (Wyszecki 1968). A clarification concerning the usage of FMC-1 and FMC-2 is given by Chickering (1971).

Allen (1971) indicates (only the abstract of his paper is available at this time) a new analysis of the PGN color-matching ellipses of MacAdam (1942). The color-matching ellipses can be expressed in terms of the variance of color

matching around certain color centers in given directions. Each of these variances is the sum of three component variances representing the uncertainties in the basic process mediated by the visual mechanism. These processes can be represented by spectral tristimulus functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$. The transformation matrix can be optimized so as to obtain the best fit between the sum of the variances calculated from the spectral tristimulus values and the experimentally determined variances.

Vos and Walraven (1970 *a* and *b*, 1971) propose a line element which is based on a zone-fluctuation theory of vision (Walraven 1966, 1968, Walraven and Bouman 1966). In accordance with this theory, color stimuli are first processed in a Helmholtz-type three-receptor zone, followed by a Hering-type neural-conversion zone in which a brightness signal and two antagonistic (red-green and yellow-blue) signals are formed. Color discrimination is assumed to be essentially limited by the photon noise. At higher luminances, color discrimination is reduced because of the impeding saturation of the discrimination processes. Many of the assumptions underlying the line element are based on experimental data gathered from various sources.

The numerical expression for the line element in its most general form is rather complex. However, it reduces to a numerically more manageable form when applied to predicting specific visual discrimination data. The prediction of the PGN ellipses (MacAdam 1942) appears to be reasonable.

Vos and Walraven also point out a close relationship between their line element and that proposed by Stiles (1946). Earlier, Trabka (1968) had given an intriguing new interpretation of Stiles' line element by deriving it as the weighted sum of a set of nonideal radiation detectors having a dark current and dead time.

Another proposal for a line element has been made by Shklover (1970) as a further development of his earlier work (Shklover 1957). Shklover postulates four stages. The color stimulus is first processed by three types of receptors with spectral characteristics linearly related to the color-matching functions of the CIE standard colorimetric observer. In the second stage, the receptor signals are attenuated in a logarithmic manner followed, in the third stage, by a conversion to two differential signals and one summation signal in accordance with two antagonistic chromatic processes (red-green and blue-green) and one brightness mediating process postulated in the underlying theory. (Note, one of the two chromatic processes is a blue-green process rather than the usually assumed yellow-blue process.) In the fourth stage of Shklover's model, essentially cortical responses are processed and signals characterizing hue, saturation, and brightness are generated.

Shklover has applied his line element to some color-discrimination data such as loci of constant hue and saturation, and to wavelength discrimination. Reasonable agreement between his predictions and the actually observed

data is indicated. No predictions of color-matching ellipses are available, such as those observed by PGN (MacAdam, 1942).

Initially, a line element (ds) is intended to predict *small* color differences of threshold or near-threshold size. When a larger difference between colors A and B is to be predicted, it is often assumed that the minimum value of the line integral, $\int_A^B ds$, between points A and B representing the two given colors, is a measure of the perceived color difference. The minimum value is obtained when the integration follows the *geodesic line* between A and B . There is some indication that larger color differences are not related to threshold-size differences in the simple way defined by the line integral. Instead, a nonlinear relation may exist.

MacAdam (1963) analyzed color scales which had been determined experimentally by the Committee on Uniform Color Scales of the Optical Society of America (Judd, 1966). The committee-scale values were derived statistically from paired-comparison judgments made on relatively large color differences, but within a restricted (x, y)-chromaticity domain. MacAdam found that the scale values $\bar{H}$ were nonlinearly related to chromaticity differences. The

$$\bar{H} = H [g_{11} (\Delta x)^2 + 2g_{12} \Delta x \Delta y + g_{22} (\Delta y)^2]^{p/2}$$

relation gave the best fit when the power p was somewhere between 0.65 and 0.80. The constant H is a convenient normalizing factor and the g_{ik} -coefficients are optimized to obtain the best linear distortion of the restricted (x, y)-chromaticity domain. The case of $p = 1$ would be the linear case often assumed in color-difference evaluations.

Sugiyama (1964) has questioned the validity of MacAdam's nonlinear hypothesis and suggested that a bias in the derivation of the committee-scale values $\bar{H}$ may be responsible for the nonlinear relationship given above. The rating scale used in the paired-comparison experiment may have been too coarse and the subsequent statistical analysis may thus not reflect the true visual judgments. The work by Sugiyama and Wright (1963) is cited to support the argument against the nonlinear hypothesis. Moreover, earlier work by Indow and Kanazawa (1960) on multi-ratio judgments of color differences has resulted in relationships suggesting a power p greater than unity. The controversy has not been resolved.

Friele (1970), suspecting that color differences of threshold size are not related to color differences well above threshold size, reconsidered the multi-stage Müller theory of color vision, he had already used in his line element, to derive a modified metric applicable to large color differences. The parameters of his model were adjusted so as to yield a best fit to the spacing of the Munsell system. The fitting is reasonably successful. The derived color-difference formula differs in many ways from the line element for threshold-size differences.

Recent reviews on line elements are included in publications by LeGrand (1970), Sugiyama (1964), Vos and Walraven (1972), and Wyszecki and Stiles (1967).

4.2 *Color-Difference Formulae*

4.2.1 *New Formulae and Modifications of Existing Formulae*

A few new color-difference formulae have been developed in recent years, and several modifications to existing formulae have been proposed.

One of the new color-difference formulae is Friele's (1970) development, discussed above, based on the Müller theory of color vision and on the color spacing in the Munsell system. No tests on the formula have been made as yet with other visual discrimination data.

Richter, K. (1969) proposes a color-difference formula based on concepts involving an opponent-colors theory of vision. The "unique hues" of color perception play an important part in the underlying assumptions of Richter's model (Richter, K., 1970, 1971). The color-difference formula is a Euclidean-type difference equation involving a lightness component and two chromatic components. The model is applied to the spacing of the colors in the Munsell system and a reasonable prediction of the spacing is obtained. The PGN color-matching ellipses (MacAdam, 1942) are not as well reproduced. Richter recommends further work on his model.

A theoretical study is made by Sugiyama (1964) suggesting a color-difference formula that conforms with hyperbolic geometry, that is a color space with negative Gaussian curvature.

Fukuda and Fujii (1965) modify the Adams-Chromatic-Value System by replacing the Munsell-Value function (a fifth-order polynomial) by the W^* -function of the CIE 1964 (U^* , V^* , W^*)-system. Color-difference calculations in the modified system are easier without loss of accuracy (see also Morton, 1970).

In this connection, another, somewhat earlier modification of the Adams-Chromatic-Value System comes to mind, which is not aimed at simplification of the calculations involved in evaluating differences, but is aimed at improving its prediction of the color spacing of the Munsell system. This modification was suggested by Taguti and Sato (1962) and involves the introduction of powers to the differencing components. The numerical expressions are rather complicated.

Sato and Hioki (1969) modified the CIE 1964 (U^* , V^* , W^*)-system so as to conform better with the color spacing defined by the Munsell-renotation system which they assume to be perfect. They propose the following modification of the CIE 1964 color-difference formula:

$$\Delta E' = f[(\Delta U')^2 + (\Delta V')^2 + n^2 (\Delta W')^2]^{\frac{1}{2}}$$

where f is a constant to make $\Delta E'$ equal to ΔE_{CIE} for Munsell colors at Munsell Value 5/. The factor n is a weighting factor dependent on observing conditions, and the new coordinates U' , V' , W' are defined by

$$\begin{aligned} W' &= 25 Y^{1/3} - 17 & 1 \leq Y \leq 100 \\ U' &= 25 \tilde{W}(u - u_0) \\ V' &= 25 \tilde{W}(v - v_0) \quad \text{with } \tilde{W} = (W')^{5/6}, \text{ where} \end{aligned}$$

W' is identical to the original W^* and the chromaticity coordinates u , v are computed as before. The modified formula reduces substantially the weight of the lightness component (W^*) relative to the chromatic components (U^* , V^*).

From an analysis of the color-matching ellipsoids of Brown (1957) and Brown and MacAdam (1949), Sugiyama (1968) deduces that the factor 25 in the equations for U' and V' given above ought to be 13, but in a subsequent note, Sugiyama (1969) changes it to 19 because of a misprint he discovered in Brown's (1957) paper.

Judd and Yonemura (1970) found that the CIE 1960 UCS diagram can be generated by the second stage of the Müller theory of color vision. The angles subtended at the convergence points of the chromaticity-confusion lines for protanopia and tritanopia by any two chromaticity points are measures of the size of the difference perceived between the two corresponding chromaticities by a protanopic and tritanopic observer, respectively. In the Müller second stage, normal color vision is assumed to be the result of combining the protanopic and tritanopic discriminative ability. A simple formula evolves for the evaluation of the differences between colors of the same luminance. The model is tested on wavelength discrimination data for protanopes, tritanopes, and normal trichromats with encouraging results.

This model has been improved further by Yonemura (1970) by optimizing some of its parameters to obtain a better fit between predictions made by the model and experimentally determined chromaticness differences (Wright, 1941) and loci of constant hue of the Munsell renotation system. A curvilinear chromaticity diagram is obtained which resembles that derived by MacAdam (1970 *b*, 1971 *a*). (See also Judd, 1971.)

MacAdam's (1970 *b*, 1971 *a*) curvilinear chromaticity diagram is based on the standard deviations of color matching (MacAdam, 1942, Brown and MacAdam, 1949, Brown, 1957). The color-matching ellipses can be converted into near circles of equal size by a nonlinear transformation of the CIE (x , y)-chromaticity coordinates to new rectangular coordinates ξ , η . The deviations of the near circles from ideal circles are considered small and non-systematic; the mean-square error of fitting the selectively averaged color-matching data for 14 observers is estimated to be approximately 24 percent.

At this point, it is worth recalling the curvilinear chromaticity diagram derived

by Farnsworth (1958), which converted the PGN-ellipses of MacAdam (1942) into near circles of equal size with a similar success. However, Farnsworth did not give a numerical expression for his diagram.

MacAdam calls his (ξ, η) -diagram a geodesic chromaticity diagram. The distance between two chromaticity points (ξ_1, η_1) and (ξ_2, η_2) , measured along the straight line connecting the points, is proportional to the difference perceived between the two (equiluminous) colors represented by the two points. When straight lines of the (ξ, η) -diagram are transformed back into the (x, y) -diagram, they generally will convert into curved lines. If, in particular, one converts a pencil of straight lines, with its apex at the point (ξ_0, η_0) representing the achromatic stimulus, a pencil of curved lines emerges from the corresponding point (x_0, y_0) in the (x, y) -chromaticity diagram. The interesting result of this is the resemblance of these curved lines with experimentally determined hue lines.

The measure for the perceived chromaticity difference ΔC of two equiluminous colors of chromaticities (ξ_1, η_1) and (ξ_2, η_2) is, in accordance with MacAdam (1970 *b*),

$$\Delta C = [(\xi_1 - \xi_2)^2 + (\eta_1 - \eta_2)^2]^{\frac{1}{2}}$$

If a luminance difference $\Delta Y = Y_1 - Y_2$ is permitted between the two given colors, MacAdam (1970 *b*) suggests the formula

$$\Delta E = [(K_1 \Delta C)^2 + (k K_2 \Delta Y/Y)^2]^{\frac{1}{2}}$$

where K_1, K_2 are functions of Y as suggested by Chickering (see Wyszecki, 1968), and k is a constant depending on application and observing conditions.

In his subsequent paper on the same subject, MacAdam (1971 *a*) suggests a somewhat different and less specific method of incorporating luminance differences in the calculation of ΔE . He suggests to simply add to $(\Delta C)^2$ the quantity $(\Delta Y)^2$, weighted by any suitable function of Y to convert ΔY to a lightness increment. MacAdam further suggests that, for some applications, it may be desirable to multiply the calculated chromaticity difference by a function of Y . But, for evaluation of small color differences, of the order of production tolerances in most industries, MacAdam believes, that this multiplication should not be applied to the coordinates themselves because the color-matching ellipsoids are almost certainly not tipped in the manner implied by this procedure. The latter point raised by MacAdam will be brought up again later on.

Nonlinear transformations of the (x, y) -chromaticity diagram, to yield more nearly perceptually uniform chromaticity diagrams, have also been proposed by Matveev (1964), Matveev and Beljaeva (1965), and Richard (1966). The work by Richard (1966) explores the application of two-opponent processes of the Müller type in the visual mechanism to the derivation of numerical expressions that would convert the Munsell renotation network of chroma

lines and hue lines at Munsell Value 5/ into an ideal network of equally spaced concentric circles (chroma) and straight lines (hue) spaced at equi-angular intervals. The theoretical results are reasonably successful.

Blottiau and Penciolelli (1966) modify the CIE 1960 UCS diagram to obtain PGN color-matching ellipses (MacAdam, 1942) in closer agreement with circles. They suggest an (s, t)-diagram which accomplishes this with some success, but the near circles they obtain differ in sizes. The argument is that circular shape is of greater importance in practice than variations in size from one location to another of the chromaticity diagram.

There is a plea by Richter, M. (1968) to study the color-difference formula based on the DIN-Color System and an announcement by McLaren and Coates (1970) that the Adams-Nickerson color-difference formula (also known as the Adams-Chromatic-Value formula) is recommended for general use in the (British) textile industry.

The Adams-Nickerson formula is described in detail by McLaren (1970*b*). This formula is said to have been found superior to the four formulae recommended by the CIE for further study. Although the announcement (McLaren and Coates, 1970) makes no reference to specific work of testing the various formulae, it is known that the correlation coefficients quoted by McLaren (1970 *c*) in his paper on 'Colour Passing—Visual or Instrumental?' have led the Colour Measurement Committee of the Society of Dyers and Colourists to recommend the Adams-Nickerson formula.

Recent reviews on color-difference formulae and related subjects have been given by Billmeyer (1969 *a, b, c*, 1970), Billmeyer and Saltzman (1966), Hemmendinger (1970), Sugiyama (1969 *a, b*, 1970 *a, b*), Judd and Wyszecki (1963), Wright (1969), Wyszecki and Stiles (1967), and McLaren (1969).

4.2.2 Numerical and Graphical Aids for Evaluating Color Differences

Tsukada et al (1966) have prepared Tables of (x, y)-chromaticity coordinates for 100 Munsell-Value planes from 1.0/ to 9.9/ at 0.1 Value intervals. A Lagrange-interpolation technique is used.

Fuwa and Sugiyama (1965) and Sugiyama and Fuwa (1966) have converted the (x, y)-chromaticity coordinates of the Munsell-renotation system to (u, v)-coordinates of the CIE 1960 UCS diagram. Graphs as well as Tables are given.

The CIE 1931 and 1964 spectral tristimulus values $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ and $\bar{x}_{10}(\lambda)$, $\bar{y}_{10}(\lambda)$, $\bar{z}_{10}(\lambda)$, respectively, have been transformed to $\bar{u}(\lambda)$, $\bar{v}(\lambda)$, $\bar{w}(\lambda)$, and $\bar{u}_{10}(\lambda)$, $\bar{v}_{10}(\lambda)$, $\bar{z}_{10}(\lambda)$, respectively, in accordance with the transformation defining the CIE 1960 USC diagram (Nimeroff 1964, superseded by CIE 1971).

Sugiyama (1964) devised an approximate technique of calculating rapidly the CIE 1964 coordinates U^* , V^* , W^* . He makes use of the derivatives dU^* , dV^* , dW^* .

A differential technique has also been used by Bridgeman (1964) to calculate

approximate color differences in the Adams-Chromatic-Value System.

Blackwood and Billmeyer (1966) have published a Fortran II computer program to calculate color differences in accordance with an earlier version of the MacAdam-modified Friele formula. A Fortran IV program for the FMC-2 formula has been published by Billmeyer and Smith (1967).

Fiorenzi and Checchi (1970/71) describe a program for calculating CIE 1964 color differences and NBS color differences on an electronic desk calculator.

Richter, M. (1964) has designed a slide rule to facilitate the determination of DIN-Dunkelstufe in the DIN color-difference formula.

Charts for the rapid evaluation of color differences in accordance with the PGN color-matching ellipses (MacAdam 1952) have been prepared by Foster (1966). There are 89 charts covering the CIE (x, y)-space. They are related to previously developed charts by Davidson and Hanlon (1955) and Simon and Goodwin (1958).

Chickering (1969) describes the construction of (x, y)-chromaticity diagrams, for different values of Y , which show contour lines of the constant partial derivatives of the CIE tristimulus values with respect to the lightness component and the two chromatic components characterizing the FMC-2 color-difference formula. Such contour diagrams may be used to estimate the significance of errors in CIE tristimulus values in terms of perceived differences between colors in the various parts of the CIE tristimulus space.

4.2.3 Numerical Intercomparison of Color-Difference Formulae

In a comprehensive review of commonly used color-difference formulae, McLaren (1969) notes the confusing situation regarding scaling factors which determine the relative sizes of color-difference units of different formulae. The question is raised whether scaling factors should be recommended to bring the size of different color-difference units in closer agreement with that of the CIE 1964 formula.

Such a procedure would not be desirable as it would tend to confuse the issue even further. Basically there can be no single factor which would adequately relate the sizes of the differences calculated by one formula with those calculated by another formula. Although the corresponding *average* differences calculated by means of two formulae for a given set of pairs of colors may be made to agree by applying a factor f , the variance between the corresponding sets of calculated color differences will generally be very large and thus make f a factor of questionable practical value. The factor will also vary with the particular set of pairs of colors used for its determination. For a discussion of the problem see also MacAdam (1969), Saltzman (1969), Richter and Weise (1970), and Schultze (1970).

The strikingly large inherent differences between different color-difference formula have been demonstrated numerically and graphically by Ganz (1966),

Schultze (1970), Richter and Weise (1970), and Steen and Tonnquist (1970). For example, Schultze (1970) calculates, at different locations in (x, y, Y) -space, loci of constant ΔE from different color-difference formulae. These loci are of ellipsoidal shape and, at a given center, their corresponding orientations, sizes, and shapes vary greatly from one another, particularly for highly saturated colors. The discrepancies between corresponding loci of constant ΔE can be used as a guide for designing specific visual tests on the validity of the color-difference formulae.

Some formulae generate ellipsoidal loci of constant ΔE which are tilted in (x, y, Y) -space against the Y -axis; that is, their highest points (x_h, y_h) , relative to the (x_0, y_0) -location of their center, is displaced toward the chromaticity point (x_n, y_n) of the achromatic stimulus. The angle of tilt increases with the distance of the center (x_0, y_0) from the achromatic point (x_n, y_n) . Formulae which generate such tilted loci are the CIE 1964 formula and those based on the Munsell system (e.g. the cube-root formula, the Munsell-Godlove formula, and the Adams-Chromatic-Value formula). The formulae based on color-matching ellipses (MacAdam, 1942), with an added lightness component, do not generate tilted loci, that is, their highest points (x_h, y_h) always lie exactly above the center (x_0, y_0) .

MacAdam (1971 *b*) has looked at the color-matching ellipsoids of Brown and MacAdam (1949) and Brown (1957) and found no significant tilts of the kind indicated by some of the formulae.

Schultze and Gall (to be published) have conducted an experiment with saturated red and blue colors involving experienced industrial color matchers. Their experiment indicates that, if there is any tilt of the ellipsoidal loci of constant ΔE at all, it must be very small. Schultze and Gall conclude that the tilt generated by the CIE 1964 formula is much too pronounced. The FMC-2 formula, the predictions by the Simon-Goodwin method, and the DIN formula are in good agreement with the requirement of no tilt as indicated by the industrial visual assessment of dyed textile samples.

An earlier experiment by Nickerson and Stultz (1944), summarized by Nickerson (1944), comes to mind. In this experiment, the visual judgments on painted fabrics favored the formulae which were based on the color spacing of the Munsell system. In particular, a set of weak reddish-brown fabrics were in good agreement with visual assessments because of the tilt of the loci of constant ΔE calculated from such Munsell-based formulae. This latter interpretation was offered by Judd in the discussion of Nickerson's (1944) paper. It must be noted that both the paper by Schultze-Gall and by Nickerson refer to visual assessments of the acceptability of given sizes of color differences. We will see later that there are strong indications that *acceptability* and *perceptibility* data do not necessarily correlate well.

Finally, Nimeroff (1971) asks whether the CIE 1964 ($U^*V^*W^*$)-space has

a spectrum locus. In noting that the spectrum locus shifts for different luminance levels and with chromaticity of the adapting source, Nimeroff considers this a serious limitation of the CIE 1964 space. Details of Nimeroff's arguments are expected to be in the paper of which only the abstract is available at this moment.

4.2.4 Accuracy of Color-Difference Formulae

A color-difference formula is considered accurate if it predicts correctly the relative magnitude of two color differences an observer with normal color vision perceives when viewing two given pairs of color stimuli.

In any meaningful determination of accuracy it is important that the observer makes color-difference judgments under conditions of observation that fall within the range of those conditions to which the given color-difference formula applies. The main parameters that define the conditions of observation are the

- i) geometrical configuration of the field of view consisting of the test stimuli and stimuli surrounding the test stimuli (sizes, shapes, separation, and relative position to one another);
- ii) luminances and relative spectral power distributions of test and surrounding stimuli.

It is equally important that the observer makes color-difference judgments in direct response to his *visual perception* which must *not* be biased by considerations of the *acceptability* of the perceived color difference for a certain application.

All color-difference formulae are intended to predict the *perceptibility* or *noticeability* of color differences. They are not intended to predict whether a color difference between a standard color and a test color will be objectionably large or whether it will be an acceptable difference. Naturally, differences between stimuli which cannot be perceived will always be acceptable. However, it is not necessarily true that a perceived difference is never accepted. In fact, in many applications a quite noticeable difference between standard and test color is accepted; and the standard and test color are then said to be a *commercial match*.

The difference between the perceptibility and acceptability of color variations from a standard color must be kept in mind whenever color-difference formulae are applied to industrial color-production-control tasks. Failure of a color-difference formula to predict accurately the *acceptability* of a color difference, for a certain application, generally cannot be interpreted as failure of the color-difference formula to predict accurately the *perceptibility* of the color difference. However, a color-difference formula can be used to describe acceptable color variations from a given standard in terms of color differences perceived under certain conditions of observation. In addition to the magnitude of the perceived color difference, some formulae also permit the evaluation of the perceived

direction in which the test color deviates from the standard, e.g. in terms of the perceived shifts in hue, saturation, and lightness.

In recent years, only a few accuracy tests have been made on color-difference formulae by comparing the predictions of the formulae with experimental data describing the perceptibility of color differences. By far the majority of tests involving color-difference formulae are tests on the applicability of these formulae to predicting the acceptability of color variations from a given standard with regard to specific industrial problems. This latter kind of test will be reviewed in the next section (4.2.5).

Wyszecki and Wright (1965) described a field trial of the CIE 1964 color-difference formula. Five sets of experimental data (three from Sugiyama and Wright (1963, 1964) and Wright, H. (1965), and two new sets) were used. The observed 176 color differences, expressed in terms of interpoint distances statistically derived from approximately 30 000 ratio- and paired-comparison judgments, were compared with the predictions made by the CIE formula. The predicted color differences ranged from 4 to 60 CIE units of color difference, with the majority between 10 and 30 CIE units, that is between medium-large to very large differences. The accuracy of predicting the observed color differences was found to be reasonable. Approximately 90% of the observed differences were predicted correctly within a boundary of $\pm 25\%$ of the calculated CIE units. The range of $\pm 25\%$ means, that in certain cases, two color differences observed to be of equal size, may be calculated to differ by 50%. Discrepancies of this magnitude were noticed in several cases, and in a few isolated cases, even larger discrepancies were found.

The significance of the discrepancies between observation and prediction must be assessed with regard to the precision of the observations as well as the accuracy of the calibration of the samples viewed by the observers. For this particular experiment, the calibration of the samples may be considered sufficiently accurate and any uncertainties in the calibration may be assumed to have a negligible effect on the relatively large color differences involved. However, the precision of the observations was not negligibly small. For the mean color difference, obtained by a group of about 10 observers, the uncertainty was estimated to be within $\pm 24\%$ of the mean. A similar uncertainty was estimated for the mean difference obtained for an individual who made several repeat observations.

Although the field trial of Wyszecki and Wright (1965) indicated a reasonable accuracy of the CIE 1964 color-difference formula for the pairs of colors examined, the field trial cannot be considered a comprehensive test. The color differences involved in the trial were mostly large differences, of the size one normally encounters in color-order systems or coarse color scales. The differences were certainly too large to be of interest to most industrial color-control laboratories who are concerned with color differences rarely exceeding 5 CIE

units. Furthermore, the color samples of most pairs had a nearly constant luminance factor, and thus the trial did not adequately test the weighting of the lightness component W^* relative to the chromaticness components U^* , V^* specified in the CIE formula. Additional tests would also be required to explore those regions of color space more thoroughly where the formula deviated from the observations by significant amounts.

Indow (1970) compared seven color-difference formulae with visual color-difference judgments in four regions of color space. The color differences were of the order of 10 CIE units. The predictions made by the formulae and the observed differences were combined and subjected to a factor analysis to explore the correlation between the eight sets of data. From the analysis, Indow deduced that the CIE 1964 formula, the Cube-Root formula, the Nickerson-Stultz formula, the FMC-2 formula, and the NBS formula make comparable predictions and represent a congenial sub-group of the whole group. The analysis did not permit a conclusive statement regarding the accuracy of any of the formulae tested.

Sugiyama (1971) made ratio comparisons on a series of pairs of colors represented in the CIE 1964 ($U^*V^*W^*$)-space by pairs of points belonging to a cube-lattice array. The interpoint distances between the colors of each pair, derived statistically from the observed ratios of color differences, were correlated with the predictions made by means of eight different color-difference formulae, which included the CIE 1964 formula, the FMC-2 formula, and the Cube-Root formula recommended by the CIE for further study. The correlation coefficients for the eight formulae ranged from 0.73 to 0.86. An interpretation of the results is expected when Sugiyama has his paper published in full.

Billmeyer et al (1971) have conducted a visual experiment using a paired comparison technique to test which one of the two versions of the Friele-MacAdam-Chickering color-difference formula, FMC-1 or FMC-2 (Chickering, 1971), agrees better with visual data. Preliminary results indicate that FMC-2 predicts color-difference judgments more accurately. The final results are promised in the paper when published in full.

As noted above, in assessing the significance of discrepancies between observed and predicted color differences, one must take into account the random uncertainties inherent in the observed as well as predicted color differences. The color stimuli are calibrated, usually in terms of CIE tristimulus values and chromaticity coordinates, and inevitably the calibrations will have errors. Generally, the size of these errors will be small and vary with the kind of stimuli calibrated and the kind of instrumentation used for the calibration. The calibration data are used to calculate color differences between the given pairs of stimuli in accordance with a given color-difference formula, and thus the calibration errors will propagate into the calculated color differences.

Illing and Balinkin (1965) have studied the *reproducibility* of calculated color

differences by calibrating seven sets of 5 colored tiles, each set exhibiting small color differences (0.2 to 4.0 NBS units) in one of seven regions of color space. Five different types of instruments in 16 different laboratories were used for the calibration. When absolute colorimetry was used (spectrophotometry on individual tiles and calculation), the color differences, expressed in NBS units, reproduced within ± 2.5 units. Differential colorimetry (2 tiles measured relative to one another) reproduced the color differences within ± 0.15 units. These results indicate that the use of differential colorimetry is the preferable method of calibration.

The question of the *accuracy* of the calibration of color stimuli was not investigated. This is a rather complex problem, involving, among other factors, the validity of the CIE standard colorimetric observer and the geometry of illumination and viewing used in the calibrating instrument. The actual observer participating in the visual assessment of color differences may deviate from the standard observer, and the geometry of the illuminating and viewing conditions under which he makes his judgments may also deviate from those employed in the instrument. The deviations between the observers and the differences in geometry can affect the accuracy of the calibrations if the colored samples exhibit pronounced differences in spectral composition and if their reflectance factors are highly sensitive to changes in the directions of illumination and viewing.

4.2.5 *Application of Color-Difference Formulae to Predicting the Acceptability of Color Variations from a Standard*

Several extensive sets of visual data on the acceptability of color variations from a standard are available, and tests have been made to see whether such data can be predicted by a color-difference formula. The results obtained by different investigators do not always lead to similar conclusions. However, a strong indication seems to evolve from most studies that satisfactory predictions of acceptable color differences cannot be made by *any* color-difference formula which is based solely on the perceptibility of color difference. It appears that parameters, other than those associated with color perception, must be used in conjunction with a color-difference formula to derive a satisfactory method for evaluating the acceptability of a given color difference. These additional (mundane) parameters are directly associated with the kind of colored material under study and its intended or anticipated application, the technical problems and costs involved in producing the material, the wishes of the customer who ordered it, and other factors.

Thurner and Walther (1970) collected extensive visual data on the acceptability of color variations from 27 standard colors in different regions of color space. A total of 500 dyed textile samples were assessed visually by 16 experienced textile colorists. The color differences between the observed pairs of

samples were generally small and of the order of 1 to 2 CIE units of color difference.

The visual judgments of the 16 colorists were averaged for each pair of samples and compared with color differences calculated from eight color-difference formulae (Adams-Nickerson-Stultz, CIE 1964, DIN, NBS, FMC-2, Simon-Goodwin, and Saunderson-Milner).

Thurner and Walther conclude from the results that none of the color-difference formulae gives an entirely satisfactory correlation with the visual assessment. However, from among the eight formulae used, the DIN-formula and the one based on the Simon-Goodwin Charts were the best. The results also show significant variations in the corresponding judgments between repeated judgments by an individual colorist. Thurner and Walther note that these variations can be similar in magnitude to those that occur between color differences calculated from different color-difference formulae for the same pair of colors.

The earlier investigation on the size of acceptable color differences by Davidson and Friede (1953) has been recalled by several workers to test color-difference formulae.

Berger and Brockes (1966) use the data by Davidson and Friede to test the usefulness of the CIE 1964 color-difference formula to predicting the size of acceptable color differences in wool-flannel dyeings. They find that the CIE formula does not make any better predictions than other formulae that have been tested against the data by Davidson and Friede. The calculation procedure involved in the CIE formula is found to be cumbersome and the charts by Simon-Goodwin are preferred.

Kuehni (1970) uses both the Davidson-Friede (1953) data and the more recent Thurner-Walther (1970) data to test four color-difference formulae (CIE 1964, Simon-Goodwin, FMC-2, and a formula based on an extension of MacAdams's (1970 *b*) curvilinear space to three dimensions (ζ , η , W)). Kuehni's correlation diagrams, between observed data (percent acceptability of color difference) and calculated color difference, show little difference between the performance of the different formulae. However, a marginally better performance of the formula based on the Simon-Goodwin Charts is indicated.

Kuehni finds that the Thurner-Walther data can be separated into two subsets. One subset refers to acceptable sizes of color differences caused by changes of *colorant strength* and the other to acceptable sizes of color differences caused by changes of *colorant shade*. When he analyses the two subsets of data separately, Kuehni notes that the colorists accept much more frequently color differences caused by changes in colorant strength. The implications of this fact are considerable and appear to confirm that certain perceived sizes of a color difference are not necessarily acceptable sizes. Thus, a color-difference formula which may predict perceived color differences perfectly, may not

at all be suitable to predict the acceptable sizes of color differences.

Strocka (1971) also uses the Davidson-Friede (1953) and Thurner-Walther (1970) data to test eight different color-difference formulae (Simon-Goodwin, FMC-2, CIE 1964, Cube-Root, Adams-Nickerson-Stultz, NBS, Hunter's L , a , b , and DIN).

The correlations between the visual data and the predictions made by the formulae were first expressed in terms of Spearman's rank-correlation coefficient and then in terms of "wrong decisions" that might be made on the basis of the predictions by color-difference formulae.

Strocka first obtains correlations between calculated color differences and those visual assessments that apply to an individual standard color. Then he correlates the data for the entire set of standard colors. Some attention is given to the weight applied to the lightness component in each of the color-difference formulae, and different weighting factors are used to see whether changing them would lead to improvements of the performance of the formulae.

Based on his different, often extensive numerical analyses, Strocka draws several interesting conclusions. He finds that none of the color-difference formulae, in their original form, makes satisfactory predictions of visually obtained data on the acceptability of color differences. The correlation between visual assessments of acceptability and calculated color differences is poor.

The performance of all formulae improve somewhat if the weight of the lightness component relative to the chromaticness component in the formulae is reduced. This indicates that lightness differences are generally more acceptable than chromaticness differences of perceptually equal size.

The large variation between corresponding judgments of the acceptability of color differences made by different observers is also noted. If the average judgment is used as reference, the sum of the deviations (wrong decisions) from the average is used to determine the percentage variation of the visual judgments. Strocka finds that for the Davidson-Friede data this variation is 17.2% and for the Thurner-Walther data it is 19.6%. In contrast to these variations in the visual data, the different color-difference formulae deviate from the average Davidson-Friede data by 12.2 to 23.3%, and from the average Thurner-Walther data by 22.9 to 27.9%.

As the main factors responsible for the failure of color-difference formulae to predict the acceptable size of color differences, Strocka gives *i*) the inherent inconsistencies in the visual assessments, *ii*) the limited accuracy of the calibration of the color samples, and possibly *iii*) the partial independence of acceptable color differences from perceived color differences.

With regard to *ii*) Strocka notes that 16.7% of the color differences involved in Davidson-Friede's experiment were below 2 MacAdam units; in Thurner-Walther's experiment they were 31.3% below that size. The uncertainties in the calculated color differences due to calibration errors are estimated to be

at least 0.5 MacAdam units. These uncertainties cannot be ignored when small color differences are involved, as in the case of the Davidson-Friede and Thurner-Walther experiments. (See also the earlier discussion in Section 4.2.4 of this review.)

With regard to *iii*), Strocka agrees with Kuehni (1970) whose contribution has been discussed above.

In addition to the eight color-difference formulae Strocka tested, he also examined a difference formula which takes little or no account of color-difference perception, that is

$$\Delta E = 1000 [(\Delta x)^2 + (\Delta y)^2 + (0.05 \Delta Y/Y)^2]^{\frac{1}{2}}$$

This formula did not perform worse than any of the eight color-difference formulae.

McLaren (1970) explores the precision of textile color matchers and finds that experienced personnel wrongly rejects about 25% of acceptable color differences. This appears to agree with the percent variations just quoted from Strocka's (1971) paper based on the work of Davidson-Friede (1953) and Thurner-Walther (1970).

McLaren also uses the Davidson-Friede data to correlate them with six color-difference formulae (CIE 1964, Adams-Chromatic Value, Cube-Root, Godlove-Munsell, Simon-Goodwin, FMC-2). The correlation coefficients he calculates are, respectively, -0.54 , -0.58 , -0.58 , -0.55 , -0.58 , -0.62 . McLaren notes that the differences between the correlation coefficients are not statistically significant. An appropriate choice between the different formulae may then be made on the basis of other considerations, such as ease of calculation.

Robinson (1969) has conducted an extensive, but limited in scope, field trial on a production set of paint samples which show small variations in color from a given blue-gray standard color. The 31 samples were ranked by 132 observers in degrees of acceptability as a production match with the standard. An attempt was then made to predict the mean visual ranking by means of the corresponding color differences calculated between each of the production samples and the standard from different color-difference formulae. Robinson concludes that the best predictions were obtained by color-difference evaluations based on the Brown-MacAdam color-matching ellipsoid. The other formulae (modified Adams-Chromatic-Value, Hunter-Scofield, CIE 1964) make poorer predictions, particularly for a group of 6 samples exhibiting relatively large differences from the standard. The Adams-Chromatic-Value formula and the CIE 1964 formula are only marginally better than the Hunter-Scofield formula.

In a series of papers, Coates et al (1968 *a* and *b*, 1969, 1970 *a* and *b*, 1971) describe various tests on color-difference formulae and their usefulness to predict the sizes of acceptable color differences. A number of conclusions were

drawn which, in the main, corroborate the results of other investigators. None of the color-difference formulae tested (CIE 1964, NBS, Hunter-Scotfield, Cube-Root, Saunderson-Milner, FMC-2, Godlove-Munsell) makes satisfactory predictions for either paint or textile samples.

Coates et al also note that photoelectric colorimetry, used to measure color-differences directly, reproduce the measurements with little variation. The average standard deviations for several instruments were 0.11, 0.00031, 0.00043 for ΔY , Δx , Δy , respectively, resulting in variations in the ΔE values of typically 0.1 CIE units. But when the same colorimeters were used to measure the tristimulus values of individual samples, large variations were found. For example, for a mat paint sample, the measured values of X varied from 14.50 to 16.69, values of Y from 17.57 to 20.30, and values of Z from 16.99 to 19.02. When measuring textile samples special care must be taken in presenting a sufficient thickness of material with a uniform and plane surface to the instrument, and a sufficient number of readings taken at different orientations to provide a reliable average reading.

Coates et al stress their finding that in general, visual assessments of the acceptability of color variations from a standard cannot readily be connected with the predictions made by color-difference formulae, because acceptability judgments are often made under conditions of observation quite different from those to which the color-difference formulae apply.

As a result of the disappointing performance of color-difference formulae, Coates et al propose an empirical formula

$$(\Delta E)^2 = 10^5 (\Delta x)^2 + 8.10^4 (\Delta y)^2 + 0.8 (\Delta Y)^2$$

which, in accordance with their experimental data, predicts the acceptability judgments better than any of the color-difference formulae they tested.

Sugiyama (1967) uses the data by Davidson (1964) to test 6 color-difference formulae including Simon-Goodwin, Adams-Nickerson, CIE 1964, NBS, and Sugiyama-Fukuda. The correlation coefficients, between the visual assessments of the acceptability of small color differences between paint samples and the corresponding color differences calculated from the formulae, range from 0.6 to 0.71. None of the formulae tested could be considered superior to any of the others.

The data by Schultze and Gall (to be published) were already discussed earlier in connection with the tilt of loci of constant color difference from a central color in (x, y, Y) -space (Section 4.2.4). For highly saturated red and blue textile dyeings, the experimentally determined acceptability data were predicted best by color-difference formulae which do not generate tilted loci of constant ΔE . The formulae FMC-2, Simon-Goodwin, and DIN perform best in this regard, whereas the CIE 1964 formula fails.

In the abstract of a forthcoming paper, Jaeckel (1971) indicates the comple-

tion of an extensive experiment on the acceptability of color variations from a standard color (dyed textiles) and the utility of color-difference formulae for predicting the visual assessments. Jaeckel examined 18 color-difference formulae against visual acceptability data and found correlation coefficients ranging from -0.79 to -0.26 . The Cube-Root formula as modified by Morton (1970) is found to give the highest correlation followed closely by the Adams Chromatic-Value formula. The use of both these formulae results in significantly fewer wrong decisions than by an assessor. Most of the other formulae tested are significantly inferior to these two formulae.

Jaeckel also notes some problem areas that arose in his experiment. These included instrument reproducibility (he used a Harrison filter colorimeter), variations of tolerance limits with assessor group, color, surface structure, direction of change from standard, and choice of independent variable for the regression analysis of the data.

Malkin (1967) has made an experiment on ceramic wall tiles. Match acceptability was the task given to some 700 observers. The tiles were calibrated on a 'Colormaster' colorimeter and the color differences between them were expressed in terms of modified Adams color-difference units, CIE 1964 units, and read off Simon-Goodwin Charts. None of the three color-difference evaluations made entirely satisfactory predictions of the visual data.

In a re-evaluation of this experiment, Dinsdale and Malkin (1971) noticed marked differences between visual assessments and predictions made by color-difference formulae, particularly for the pink tiles. CIE 1964, Simon-Goodwin, and the FMC evaluations all give greater weight to saturation differences than to hue differences, in direct contradiction to the visual assessments.

Mudd (1964) and Mudd and Woods (1970) have studied the problem of color tolerance specifications on leather materials and examined various color-difference formulae with regard to their usefulness to the problem. They find that the CIE 1964 and the Adams Chromatic Value formulae are the best of the color-difference formulae they tested, but even these fail badly in some cases (8 or 9 times out of 30). A method of assessing color tolerances which is found to be superior to the use of color-difference formulae is based on an empirical formula involving differences in chromaticity coordinates (x , y) and luminance factor, Y .

Simon (1971) conducted a visual experiment on a number of dyed textile swatches to test the predictions made by the color-difference formulae FMC-1, FMC-2, (ξ , η) of MacAdam (1971), and those made by the Simon-Goodwin graphical method. In the abstract of his yet unpublished paper, Simon states that in most cases the visual data tended to be predicted better by the (ξ , η)-formula than by the other methods. He also favors the (ξ , η)-formula because of its convenience of use.

Schultze (1971) proposes a new correlation coefficient to assess the agreement

between acceptable sizes of color differences, determined experimentally, and corresponding perceived color differences, determined by means of color difference formulae. The correlation coefficient is defined by

$$V = \left\{ \frac{1}{n} \sum_{i=1}^n \frac{(\Delta E_A - F \Delta E_B)^2}{(\Delta E_A)^2 F^2 (\Delta E_B)^2} \right\}^{\frac{1}{2}}$$

with the normalizing factor F given by

$$F = \left\{ \frac{\sum_{i=1}^n [(\Delta E_A) / (\Delta E_B)]}{\sum_{i=1}^n [(\Delta E_B) / (\Delta E_A)]} \right\}^{\frac{1}{2}}$$

The quantities (ΔE_A) and (ΔE_B) denote the color differences of a given color pair(i) derived from visual assessments of the acceptability of color differences and from calculation by means of a color difference formula, respectively. The same coefficient, V , can also be used to intercompare the predictions made by two color-difference formulae A and B (Schultze 1970). The smaller the coefficient, the higher the correlation between the two sets of data. Perfect correlation is indicated for $V = 0$.

Schultze calculates V to intercompare each of several color-difference formulae with each of several sets of experimental acceptability data reported by different investigators. In particular, he draws from the experimental work reported by Nickerson-Stultz (1944), Davidson-Friede (1953), Coates-Warburton (1968), Coates-Day-Rigg (1969), Robinson (1969), Thurner-Walther (1970), and Schultze-Gall (1971).

The numerical material gathered by Schultze is extensive. Yet, Schultze notes that most of the available experimental data show inherent inconsistencies which handicap the analysis. The colors studied represent color space only in the region of less saturated colors, more saturated colors would be desirable.

Despite these difficulties, Schultze's analysis provides some very valuable confirmations of the results of other investigators and of the expectations expressed by many colorimetrists. It appears unlikely that any color-difference formula in its present form can predict satisfactorily the acceptability of color variations from a standard color for all applications. The many tables of V -correlation coefficients Schultze provides, reveal the occasional success of the different color-difference formulae in different applications and for different colors. However, very often the values of V become large, above 30, indicating unsatisfactory performance of the particular formula.

Schultze also comments briefly on the use of empirical formulae, such as those suggested by Strocka (1971), Coates et al (1969), and Mudd and Woods (1970). These formulae can provide satisfactory predictions for specific and limited applications, but generally will fail badly when applied to cases for which they were not optimized.

Pitt (1967) has studied the fidelity of color reproduction in photographic

prints as compared to original colored charts. The visual task involved in this experiment is basically different from those involved in the acceptability judgments made, for example, in the textile industry. Pitt found the CIE 1964 color-difference formula to agree reasonably well with visual scores. A reduction of the weighting factor of the lightness component relative to the chromaticness component in the formula improved the correlation. He also suggested that possibly hue differences should be weighted differently from saturation differences.

There are a number of miscellaneous contributions related to tolerance specifications (Kuehni, 1970, Morton, 1970, Suga, 1965, Bell, 1969, Shesternina and Alyavdin, 1968, Liebert, 1968, Staes and Verbrugghe, 1970, and Blottiau, 1964).

Review articles on color tolerance specifications have been published by Best (1968), Davidson, H.R. (1970), Hemmendinger (1970), Hemmendinger et al (1970), Ishak (1971), Johnston (1970), Jonckheere and Jacquemart (1970), MacAdam (1965), Marshall and Tough (1968).

5. STUDIES RELATED TO COLOR METRICS

5.1 *General Color Limens*

New sets of color-matching ellipses have been determined for three observers (Wyszecki and Fielder, 1971 *a*). These ellipses compare well with those published previously by Brown-MacAdam (1949) and Brown (1957), but show systematic deviations from those obtained by MacAdam's (1942) observer PGN.

Extensive data have also been obtained on matching color differences between two given color stimuli and a variable third color stimulus, presented to the observer in a triangular configuration of three juxtaposed photometric fields in a white surround (Wyszecki and Fielder, 1971 *b*; Wyszecki, 1965). The precision of setting the third color, such that the three differences between the three colors are perceived to be of equal size, can be characterized by an ellipsoid in (x, y, Y) -space, centered at the mean observed color. This ellipsoid resembles that obtained by the same observer in a direct color-matching experiment (Wyszecki and Fielder 1971 *a*). The shape and orientation of the elliptical cross-section of the color-difference-matching ellipsoid correlate well with the shape and orientation of the corresponding elliptical cross-section of the color-matching ellipsoid. However, the sizes differ; the color-matching ellipsoid being the smaller one.

Parra (1967, 1970) has measured color-discrimination loci around several central colors by determining the threshold distances from the central color in different directions. The loci appear very irregular in shape when plotted in the (x, y) -chromaticity diagram. Their shape is dramatically different from the commonly accepted elliptical shape and Parra's results are considered puzzling.

Parra (1971) is conducting further experiments with improved instrumentation.

Crawford (1970, 1971) also finds color-discrimination loci that are not elliptical but oval shaped. They tend to support Parra's results, but Crawford's loci were smooth and did not show any jaggedness.

Clarke (1967) reported on a preliminary investigation on chromaticity discrimination for different field sizes ($60'$ to $2\frac{1}{2}'$), retinal illuminations (20 to 3000 photopic trolands), and background colors. Small-field color discrimination tends towards tritanopia as field size is decreased. Clarke derives a family of uniform chromaticity charts to cover any field size up to the 2° standard field represented by the CIE 1960 UCS diagram.

A paper by Yonemura and Kasuya (1969) tackles a similar problem. Empirical formulae are derived to predict the experimental results. Related, but somewhat more restricted in application is the contribution by Judd and Yonemura (1969) on the target conspicuity and its dependence on color and angular subtense for gray and foliage green surrounds. The CIE 1964 color-difference formula is extended by introducing three new factors to account for the special observing conditions.

Connors and Siegel (1964) determined chromaticity limens along a (purple) line in the (x, y) -diagram generated by additive mixtures of the monochromatic stimuli 440 and 645 nm at a low luminance level (approx. 0.7 cd/m^2). The limens are expressed in terms of standard deviations in color matching. The standard deviation at the red end of the line is 200 times larger than the one at the blue end. Some agreement is observed with Wright's (1941) limens, but little or no agreement with those that may be deduced from the color-matching ellipses reported by MacAdam (1942).

Holtzman and Valberg (1969) (also Valberg and Holtzman, 1971) measured just-noticeable differences of 'optimal colors'. Optimal colors are stimuli whose relative spectral compositions are essentially unity in some part of the spectrum and zero in the remaining part, with no more than 2 transitions from zero to one within the visible spectrum. The minimum values of the just-noticeable differences were correlated with the 'unique hues' of color perception.

Luria and Weissman (1965) and Weissman and Kinney (1965) studied chromaticity limens along two sets of chromaticities that can be produced by red-green and by yellow-blue additive mixtures. The size of the limens is affected by the luminance level, retinal position, and exposure time.

The effects of chromatic adaptation on color discrimination have been considered by Terstiege (1971).

5.2 Wavelength Discrimination

The interest in wavelength discrimination continues to be high and a score of studies has been made. Siegel (1964, 1965, 1969 *a* and *b*) and Siegel and Siegel (1970, 1971 *a* and *b*) appear to have been most productive in this field. They

have measured wavelength discrimination by various techniques (forced choice, color naming, color scaling) and under a variety of observing conditions (different luminance levels of test and surround stimuli and exposure time). Similar studies have been made by Connors (1964), Boynton et al (1964), Krantz (1967), Shirley (1966), Jacobs and Gaylord (1967), Beare and Siegel (1967).

Pokorny and Smith (1970) superimpose stimuli of different wavelengths on the two halves of the bipartite photometric field in which the wavelength discrimination is determined. The added stimuli change the discrimination ability of the observer. The results are difficult to interpret in terms of existing discrimination data.

Kambe (1971) has studied the effects of the flicker frequency on wavelength discrimination and suggests that the yellow-blue component of the visual mechanism is affected differently than the yellow-green.

Indow and Takagi (1968) have intercompared the wavelength-discrimination data of various workers with theoretically obtained threshold data and find essential agreement.

Ruddock (1966) makes an attempt at explaining the mechanism that mediates changes in the wavelength discrimination due to changes in field size and exposure time. Integration processes in color vision are postulated.

Dimmick (1966) constructs a color-specification system on wavelength discrimination data and the concept of the 'unique hues'. The system appears to have problems with coping adequately with metameric stimuli.

5.3 *Lightness and Brightness Discrimination*

The study of brightness and lightness, and brightness and lightness scales has found continuous attention and several important contributions have been made.

Takasaki (1966) has made extensive experiments to determine lightness changes of gray samples caused by changes in the reflectance of a gray surrounding. He developed a three-constant empirical formula to predict the gray-contrast effect and to account for the 'crispness effect': the sample lightness changes rapidly with reflectance when the sample reflectance is close to that of the surround.

Related studies are those by Reuther (1966), who also examines color samples in color surrounds, by Kaneko (1964), Warren and Warren (1966) and Warren and Poulton (1966).

Bodmann (1966) and Adrian (1969) propose numerical expressions for the threshold sensitivity regarding achromatic stimuli viewed in achromatic surroundings for a wide range of luminances and different field sizes.

Hioki and Sato (1965) compare different formulae for the lightness index (including the W^* -formula of the CIE) with the 5th-power Munsell-Value function.

Jameson and Hurvich (1964) and Jameson (1970) put perceived brightness and color contrast into the framework of an 'induced response theory of vision'. A general formula relating the luminance (L) of the stimulus with the perception of brightness (B) is given, which is found to account for the experimental results; they propose

$$B = kL^n - i$$

where $k = \text{const.}$, $n = \text{const.}$, and i denotes a certain constant fraction of the brightness response equivalent to the physiological noise in the visual mechanism. Flock (1970) has presented some arguments against the Jameson-Hurvich relationship between luminance and brightness.

Bartleson and Breneman (1967) have made extensive experiments on the brightness perception in complex fields such as photographic reproductions viewed with both illuminated and dark surrounds. The resulting brightness versus luminance functions are not simple power functions; they are non-linear in log-log coordinates and of the form

$$\log B = \alpha + \beta \log L - [\gamma \exp(\delta \log L)]$$

where B = brightness estimate, L = luminance (in mL), and $\alpha, \beta, \gamma, \delta$ are parametric constants.

More data on the lightness or brightness of *chromatic* stimuli relative to *achromatic* stimuli have been reported by Kowaliski (1966, 1967, 1970, 1971) and Wyszecki (1967).

Some miscellaneous studies related to brightness are those of Cornsweet and Teller (1965), Müller (1969), Thielert and Schliemann (1970), and Harvey (1970).

Marsden (1969) has given a comprehensive review on the subject which includes additional references.

5.4 Color Perception and Color-Order Systems

In a theoretical thesis, Evans (1964) discusses the number of independent variables which are necessary to describe color perceptions produced by light sources, illumination, and both reflecting and transmitting objects. Evans believes that more than 3 (probably 4 to 6) independent perceptual variables are necessary.

In an experimental study, Evans and Swenholt (1967) find a measureable threshold between colors that appear to contain grey and colors that appear fluorescent. A functional relationship between the grey content of colors and purity is derived. The study is extended in two subsequent papers (Evans and Swenholt 1968, 1969).

Scaling hue and saturation has been investigated by Indow and Stevens (1966). Saturation is found to be related to purity and luminance, and increases as a

power function of colorimetric purity. The power depends upon the luminance level for yellow but not for other hues. Saturation as well as hue are 'photometric continua'. Related articles are those by Panek and Stevens, 1966, Indow, 1967, Warren, 1967, Onley and Klingberg, 1970.

Estimates of saturation are affected by chromatic adaptation (Jacobs, 1967).

The use of the 'minimally-distinct border' as a criterion for estimating the saturation of monochromatic stimuli gives results similar to those obtained by other techniques (Boynton and Wagner, 1971).

Tonnquist (1966) compares 'hue circles' (Ostwald, DIN, Munsell, and the Swedish 'Natural Colour System') and distinguishes between a 'symmetrical circle', divided by the 'unique hues' into 4 equal quadrants, and an 'equispaced circle', hue steps all of the same perceptual size.

Tsukuda and Minato (1965) have made a new experimental determination of a 24-step hue circle (Chiba hue circle) which agrees with the DIN-hue circle.

Robertson (1970) has presented some preliminary experimental results on lines of constant hue in the (x, y) -chromaticity diagram.

Boynton and Gordon (1965) have measured the Bezold-Brücke hue shift by a color-naming technique and compared the data with those obtained by different methods. Similar results were found.

Hård (1970) discusses the general concepts involved in color-order systems based on color perception. Steen (1970) describes experiments on scaling the perceptual attributes of 'yellowness', 'redness', 'blueness', 'greenness', 'whiteness', and 'blackness' as related to the Swedish 'Natural Color System' (Hård, 1966 *a* and *b*). Ishak et al (1970) make subjective estimates of hue, saturation, and lightness of Munsell color chips in different color surroundings. The results indicate that the method they use is a useful one and has potential applications to measuring color appearance generally.

The perceptually uniform color solid and its geometry is discussed in general terms by Jäkel-Hartenstein (1964).

Judd (1966) gives a progress report on the work of the Committee on Uniform Color Scales of the Optical Society of America.

Various color-order systems are intercompared by Kornerup (1964) and Weise (1966). Berger and Brockes (1964) have made a visual test of the uniformity of spacing in the DIN Color Charts. Their 25 observers find the hue and saturation spacing not to be uniform.

The measurement of color appearance in complex fields of view has just begun to be explored. Recent contributions to this rather complex problem are those made by Hunt (1965), Gibson (1967), and Pearson et al (1969); see also Judd (1966).

5.5 *Methodology and Statistical Analysis*

Different scaling techniques can be used to establish uniform scales of the

perception of color and color difference. The experimental results can be analyzed by different statistical methods. Different experimental techniques and different statistical methods applied to a similar visual task do not necessarily give similar results. It is of great importance to consider the basic concepts involved when conducting scaling experiments and attempting to interpret the results on the basis of a statistical analysis.

Several papers related to color metrics are of interest in this connection (Bartleson, 1970, Burnham et al, 1970, Curtis, 1970, Helm, 1964, Indow and Matsushima, 1969, Indow, 1971, Judd, 1967, Sugiyama and Wright, 1964, Sugiyama, 1966, Wright, 1965).

6. *Concluding Remarks*

Color differences are assessed either in terms of their perceptual sizes or in terms of their acceptability with regard to a specific application. The distinction between the perceptibility and acceptability of a given color difference is of fundamental importance.

Line elements and color-difference formulae are intended to predict the perceptual size of the color difference between two given stimuli in terms of some convenient unit of measure. They can be applied to express the sizes of acceptable (or tolerable) color variations from a given color standard in terms of perceptual sizes. However, generally, a locus of constant perceptual color difference from a given color standard does not coincide with a locus of constant acceptable color difference from the same standard.

The parameters of a line element and a color-difference formula are functions of the observing conditions. Normally, the parameters apply to only one specific set of conditions of observation and tests on the line element or color-difference formula are expected to be successful only if the specific conditions of observations are heeded.

In testing line elements and color-difference formulae it is important to use color stimuli that can be calibrated accurately.

Line elements are essentially intended to predict color differences of threshold size or near-threshold size. Their application to larger differences is not fully understood. Similarly, color-difference formulae essentially are intended to predict medium-sized color differences, and their application to threshold-size differences may not be justified.

Future work may move in two directions. In one direction, work on line elements, based on the functioning of the visual mechanism, may offer the greatest rewards. The application of line elements to measuring larger differences would also be of interest. The ultimate solution would be a single line element capable of predicting color differences of various sizes. With appropriate changes of its parameters the effects of different conditions of observation can be handled by the same line element.

In the other direction, practical methods may be developed which predict the acceptability of color variations from given color standards. No single method is expected to be useful to all industrial applications. However, acceptable sizes of color variation may be expressed advantageously in terms of perceived threshold sizes derived from the line element.

REFERENCES

- Adam, M. (1969), Aufgaben und Ziele einer Mathematisierung der Farbwahrnehmung. Proc. 1st AIC congress "Color 69", Stockholm, 437-442
- Adrian, W. (1969), Die Unterschiedsempfindlichkeit des Auges und die Möglichkeit ihrer Berechnung. *Lichttechnik* 21, 2A-7A
- Allen, E. (1971), A new analysis of the MacAdam ellipses. To be published
- Bartleson, C. J. (1969), Differences and ratios in color scaling. Proc. 1st AIC congress "Color 69", Stockholm, 386-390
- Bartleson, C. J., and Breneman, E. J. (1967), Brightness perception in complex fields. *J. Opt. Soc. Amer.* 57, 953-957
- Beare, A. C., and Siegel, M. H. (1967), Color name as a function of wavelength and instruction. *Percept. Psychophys.* 2, 521-527
- Bell, J. R. (1969), Direct interpretation of colorimetric readings in terms of color tolerance. *J. Soc. Dyers Col.* 88, 108
- Berger, A., and Brockes, A. (1964), Prüfung der Gleichabständigkeit der Beiblätter zur Vornorm DIN 6164 (DIN-Farbenkarte). *Farbe* 13, 167-174
- Berger, A., and Brockes, A. (1965), Prüfung der von der CIE 1963 empfohlenen Farbdifferenzformel. Proc. Int. Colour Meeting, Lucern, 331-336. *Farbe*, 14, 269-274
- Best, R. P. (1968), Colour Measurement and Colour Tolerance in the Plastics Industry. *J. Soc. Dyers Col.* 84, 133-136
- Billmeyer, F. W., and Saltzman, M. (1966), Principles of Color Technology. New York, Wiley, Interscience
- Billmeyer, F. W., and Smith, R. (1967), Optimized equations for MacAdam color differences. *Color Eng.* 5, 6: 28-29
- Billmeyer, F. W. (1969), Optical Aspects of Color, Part XII: Color spaces and color difference equations. *Opt. Spectra* 3, 3: 85-88
- Billmeyer, F. W. (1969), Optical Aspects of Color, Part XIII: New color difference equations and their relation to curved spaces. *Opt. Spectra* 3, 4: 82-85
- Billmeyer, F. W. (1969), Optical Aspects of Color, Part XV: The MacAdam color-difference metrics. *Opt. Spectra* 3, 6: 64-70
- Billmeyer, F. W. (1970), Optical Aspects of Color, Part XVI: Appropriate use of color-difference. *Opt. Spectra* 4, 2: 63-66
- Billmeyer, Jr., F. W., Campbell, E. D., Kandel G. N., and MacMillan, J. (1971), Small and moderate color differences; I. Visual comparison of FMC 1 and FMC 2 metrics. This symposium
- Blackwood, N. K., and Billmeyer, Jr., F. W. (1966), A computer program for MacAdam "PQS" color differences. *Color Eng.* 4, 2: 24-25
- Blottiau, F. (1964), Représentation photo-colorimétrique approximativement uniforme. *Rev. Opt.* 43, 126
- Blottiau, F., and Penciolelli, G. (1966), Sur les transformations homographiques du diagramme des chromaticités. *Rev. Opt.* 45, 118
- Bodmann, H. W. (1965), Der Helligkeitsunterschied als Funktion von Objekt- und Umfeldleuchtdichte. Proc. Int. Colour Meeting, Lucern, 203-207. *Farbe*, 14, 354-358
- Boynton, R. M., and Gordon J., (1965), Bezold-Brücke hue shift measured by color-naming technique. *J. Opt. Soc. Amer.* 55, 78-86
- Boynton, R. M., Schäfer, W., and Neun, M. E. (1964), Hue-Wavelength relation measured

- by color-naming method for three retinal locations. *Science* 146, 666
- Boynton M., and Wagnèr, G. (1971), Color differences assessed by the minimally-distinct border method. This symposium
- Bridgeman, T. (1964), A relation between colour differences and tristimulus values, *Farbe* 13, 162-166
- Brown, W. R. J. (1957), Color discrimination of twelve observers, *J. Opt. Soc. Amer.* 47, 137-143
- Brown, W. R. J., and MacAdam, D. L. (1949), Visual sensitivities to combined chromaticity and luminance differences. *J. Opt. Soc. Amer.* 39, 808-834
- Burnham, R. W., Onley, J. W., and Witzel R. F. (1970), Exploratory investigation of perceptual color scaling. *J. Opt. Soc. Amer.* 60, 1410-1420
- Chickering, K. D. (1967), Optimization of the MacAdam-modified 1965 Friele color-difference formula. *J. Opt. Soc. Amer.* 57, 537-541.
- Chickering, K. D. (1969), Perceptual significance of the differences between CIE tristimulus values. *J. Opt. Soc. Amer.* 59, 986-90
- Chickering, K. D. (1971), FMC color-difference formulas: Clarification concerning usage. *J. Opt. Soc. Amer.* 61, 118-122
- CIE, Colorimetry, Official Recommendations of the CIE (1971), Publication No. 15 (E-1.3.1) Paris, CIE
- Clarke, F. J. J. (1967), Effect of field-element size on chromaticity discrimination. Proc. Symposium on Colour Measurement in Industry, The Colour Group (Great Britain), 132
- Coates, E., Day, S., Provost, J., and Rigg, B. (1971), Colour difference equations for setting industrial colour tolerances. This symposium
- Coates, E., Day, S., and Rigg, B. (1969), Colour-difference measurements in relation to visual assessments. Some further observations. *J. Soc. Dyers Col.* 85, 312-318
- Coates, E., Day, S., and Rigg, B. (1969), A reassessment of some colour difference equations. Proc. 1st AIC congress "Color 69", Stockholm, 641-648
- Coates, E., Provost, J. R., and Rigg, B. (1970), Colour-difference measurements in relation to visual assessments—Effect of viewing conditions. *J. Soc. Dyers Col.* 86, 402
- Coates, E., and Rigg, B. (1968), Introduction to instrumental measurement of colour in relation to colour tolerance. *J. Soc. Dyers Col.* 84, 462
- Coates, E., and Warburton, F. L. (1968), Colour-difference measurement in relation to visual assessment in the textile field. *J. Soc. Dyers Col.* 84, 467-474
- Connors, M. M. (1964), Effect of surround and stimulus luminance on the discrimination of hue. *J. Opt. Soc. Amer.* 54, 693-695
- Connors, M. M., and Siegel, M. H. (1964), Differential color sensitivity in the purple region. *J. Opt. Soc. Amer.* 54, 1374-1377
- Cornsweet, T. N. (1970), Visual Perception. New York, Academic Press.
- Cornsweet, T. N., and Teller, D. Y. (1965), Relation of increment thresholds to brightness and luminance. *J. Opt. Soc. Amer.* 55, 1303-1308
- Crawford, B. H. (1969), Just perceptible chromaticity shifts. Proc. 1st congress "Color 69", Stockholm, 302-311
- Crawford, B. H. (1971), Brightness units and threshold units. This symposium
- Curtis, D. W. (1970), Magnitude estimations and category judgments of brightness and brightness intervals: a two-stage interpretation. *J. Exp. Psychol.* 83, 201-208
- Davidson, H. R., and Friede, E. (1953), The size of acceptable color differences. *J. Opt. Soc. Amer.* 43, 581-589
- Davidson, H. R., and Hanlon, J. J. (1955), Use of charts for rapid calculation of color difference. *J. Opt. Soc. Amer.* 45, 617-20
- Davidson, S. L. (1964), So you want to set color tolerances. Official Digest (Federation of Societies for Paint Technology) 38, 311
- Dimmick, F. L. (1965), Color specification based on just noticeable differences of hue. *Color Eng.* 4, 1: 20-26 Also: Proc. Int. Colour Meeting, Lucern 321-330. *Farbe*, 15, (1966), 261-270

- Dinsdale, A., and Malkin, F. (1971), Colour tolerance of ceramic walltiles. This symposium
- Evans, R. M. (1964), Variables of perceived color. *J. Opt. Soc. Amer.* 54, 1467-1474
- Evans, R. M., and Swenholz, B. K. (1967), Chromatic strength of colors: Dominant wavelength and purity. *J. Opt. Soc. Amer.* 57, 1319-1324
- Evans, R. M., and Swenholz, B. K. (1968), Chromatic strength of colors. II. The Munsell system. *J. Opt. Soc. Amer.* 58, 580-584
- Evans, R. M., and Swenholz, B. K. (1969), Chromatic strength of colors. III. Chromatic surrounds and discussion. *J. Opt. Soc. Amer.* 59, 628-634
- Farnsworth, D. (1957), A temporal factor in colour discrimination. *Proc. Symposium on Visual Problems of Colour*, Teddington, 431-444; Also published by Chemical Publishing Co., New York, 1961 (there, see Vol. 2, p. 65).
- Fiorenzi, G., and Checchi, R. (1970/71), Aufstellung eines Rechenprogramms zur Bestimmung der Differenz zweier Farben. *Mikrochimica Acta*, 176
- Flock, H. R. (1970), Jameson and Hurvich's theory of brightness contrast. *Percept. Psychophys.* 8, 118-124
- Foster, R. S. (1966), A new simplified system of charts for rapid color difference calculations. *Color Eng.* 4, 1: 17-19
- Friele, L. F. C. (1961), Analysis of the Brown and Brown-MacAdam colour discrimination data. *Farbe* 10, 193-224
- Friele, L. F. C. (1965), Further analysis of color discrimination data. *J. Opt. Soc. Amer.* 55, 1314-1319
- Friele, L. F. C. (1966), Friele approximations for color metric coefficients. *J. Opt. Soc. Amer.* 56, 259-260
- Friele, L. F. C. (1965), Further analysis of colour discrimination data. *Proc. Int. Colour Meeting*, Lucern, 302-314. *Farbe*, 14, 192-204
- Friele, L. F. C. (1967), Development of a metric for colours and colour differences. *Proc. Symposium on Colour Measurement in Industry*, The Colour Group (Great Britain), p. 243
- Friele, L. F. C. (1969), Preliminary analysis of the Munsell colour system in terms of the Müller theory. *Proc. 1st AIC congress "Color 69"*, Stockholm, 275-290
- Fukuda, T., and Fujii, K. (1965), A modified chromatic-value system. *Acta Chromatica* 1, 200
- Fuwa, M., and Sugiyama, Y. (1965), Munsell rennotations on the MacAdam (u , v) diagram. *Acta Chromatica* 1, 199
- Ganz, E. (1965), A comparison of the Adams-Nickerson, DeKleins, 1960 CIE-UCS and the Wyszecki colour co-ordinate systems. *Proc. Int. Colour Meeting*, Lucern, 331-414
- Farbe*, 14, 296-301
- Gibson, N. (1967), The measurement of subjective colour, *Opt. Acta* 14, 219-243
- Graham, C. H. (1965), *Vision and Visual Perception*. New York, Wiley
- Hård, A. (1965), Philosophy of the Hering-Johansson Natural Colour System. *Proc. Int. Colour Meeting*, Lucern, 357-366. *Farbe*, 15, (1966), 287-295
- Hård, A. (1965), A new colour atlas based on the Natural Colour System by Hering-Johansson. *Proc. Int. Colour Meeting*, Lucern, 367-375. *Farbe*, 15, (1966), 296-304
- Hård, A. (1969), Qualitative attributes of colour perception. *Proc. 1st AIC congress "Color 69"*, Stockholm, 351-368
- Harvey, Jr., L. O. (1970), Flicker sensitivity and apparent brightness as a function of surround luminance. *J. Opt. Soc. Amer.* 60, 860-864
- Helm, C. E. (1964), Multidimensional ratio scaling analysis of perceived color relations. *J. Opt. Soc. Amer.* 54, 256-262
- Hemmendinger, H. (1970), Development of color difference formulas. *J. Paint Technol.* 42, 132
- Hemmendinger, H., Davidson, H. R., and Johnston, R. M. (1970), Color differences and tolerances. *J. Paint Technol.* 42, 151
- Hioki, R., and Sato, M. (1965), Comparison of formulas for the Munsell value scale with respect to visually uniform spacing. *Acta Chromatica* 1, 174
- Holtzmark, T., and Valberg, A. (1969), Colour discrimination and hue. *Nature* 224, 366-367

- Hunt, R. W. G. (1965), Measurement of color appearance. *J. Opt. Soc. Amer.* 55, 1540-1551
- Hunt, R. W. G. (1967), *The Reproduction of Colour*. New York, Wiley
- Illing, A. M., and Balinkin, I. A. (1965), Precision in measurement of small color differences. *Bull. Am. Ceramic Soc.* 44, 956
- Indow, T. (1967), Saturation scales for red. *Vision Res.* 7, 481-495
- Indow, T. (1969), The uniformity among eight measures for color differences, *Proc. 1st AIC congress "Color 69"*, Stockholm, 664-670
- Indow, T., and Kanazawa, K. (1960), Multidimensional mapping of Munsell colors varying in hue, chroma, and value. *J. Exp. Psychol.* 59, 330-336
- Indow, T., and Matsushima, K. (1969), Local multidimensional mapping of Munsell color space. *Acta Chromatica* 2, 16-24
- Indow, T., and Ohsumi, K. (1971), Multidimensional mapping of sixty Munsell colors through nonmetric procedure, *This symposium*
- Indow, T., and Takagi, C. (1968), Hue-discrimination thresholds and hue-coefficients. Comparisons among data. *Jap. Psychol. Res.* 10, 179-190
- Indow, T., and Stevens, S. S. (1966), Scaling of saturation and hue. *Percept. Psychophys.* 1, 253-272
- Ishak, I. G. H., Bouma, H., and Bussel, H. J. J. van (1970), Subjective estimates of colour attributes for surface colours. *Vision Res.* 10, 489-500
- Ishak, I. G. H., and Roylance S. (1971), Colour tolerances in the paint industry. *This symposium*
- Jacobs, G. H. (1967), Saturation estimates and chromatic adaptation. *Percept. Psychophys.* 2, 271-274
- Jacobs, G. H., and Gaylord, H. A. (1967), Effects of chromatic adaptation on color-naming. *Vision Res.* 7, 645-653
- Jaeckel, S. M. (1971), The utility of colour-difference formulae for match-acceptability decisions. *This symposium*
- Jäkel-Hartenstein, B. (1964), Empfindungsgemässe Farbkörper. *Farbe* 13, 201-207
- Jameson, D. (1969), Brightness scales and their interpretation. *Proc. 1st AIC congress "Color 69"*, Stockholm, 377-385
- Jameson, D., and Hurvich, L. M. (1964), Theory of brightness and color contrast in human vision. *Vision Res.* 4, 135-154
- Johnston, R. (1970), Applications for color difference formulas. *J. Paint Technol.* 42, No. 542, 159
- Jonckheere, J., and Jacquemart, J. (1969), Enquête sur les tolérances pratiques de couleur dans l'industrie française. *Proc. 1st AIC congress "Color 69"*, Stockholm, 734-744
- Judd, D. B. (1965), Color appearance. *Proc. Int. Colour Meeting, Lucern*, 27-51. *Farbe* 14, 2-26
- Judd, D. B. (1965), Progress report for O.S.A. Committee on uniform color scales. *Proc. Int. Colour Meeting, Lucern*, 399-407. *Farbe*, 14, 287-295
- Judd, D. B. (1967), Interval scales, ratio scales, and additive scales for the sizes of differences perceived between members of a geodesic series of colors. *J. Opt. Soc. Amer.* 57, 380-386
- Judd, D. B. (1968-'69), Ideal color space. *Palette*, Nos. 29, 30, 31 Also: *Color Eng.* 8 2: 37-52 (1970).
- Judd, D. B. (1971), Perceptually uniform spacing of equiluminous colors and the loci of constant hue: *This symposium*
- Judd, D. B. and Wyszecki, G. (1963), *Color in Business, Science and Industry* (2nd ed.). New York, Wiley
- Judd, D. B., and Yonemura, G. T. (1969), Target conspicuity and its dependence on color and angular subtense for gray and foliage green surrounds. *NBS Report* 10120
- Judd, D. B., and Yonemura, G. T. (1969), CIE 1960 UCS diagram and the Müller theory of color vision. *Proc. 1st AIC congress "Color 69"*, Stockholm, 266-274 Also: *J. Res. Nat. Bur. Stds.* 74A, 23
- Kambe, N. (1971), Wavelength discrimination measured by sinusoidally alternating colored stimuli. *This symposium*
- Kaneko, T. (1964), A reconsideration of the Cobb-Judd lightness function. *Acta Chromatica*, 1, 103 (1964)

- Kornerup, A. (1964), Ein Vergleich der Farbton-Teilungen der drei skandinavischen Farbsysteme von Hesselgren, Barding und Kornerup-Wanscher mit der des Farbsystems DIN 6164. *Farbe* 13, 23-46
- Kowaliski, P. (1965), Etalonnage de l'espace des couleurs en luminosité apparente. *Proc. Int. Colour Meeting, Lucern*, 344-350. *Farbe*, 14, 302-308
- Kowaliski, P. (1967), Le problème de la luminance équivalente. *Rev. Opt.* 46, 359-369
- Kowaliski, P. (1969), Luminances équivalentes élevées des couleurs. *Proc. 1st AIC congress "Color 69"*, Stockholm, 418-426
- Kowaliski, P. (1971), Equivalent luminances and the reproduction of colors. This symposium
- Krantz, D. H. (1967), Small-step and large-step color differences for monochromatic stimuli of constant brightness. *J. Opt. Soc. Amer.* 57, 1304-1316
- Kuehni, R. (1970), A practical interpretation of calculated small color differences. *J. Amer. Assoc. Text. Chem. and Colorists*, 2, 404
- Kuehni, R. (1970), The relationship between acceptability and calculated color differences on textiles. *Color Eng.* 8, 1: 47-53
- LeGrand, Y. (1969), Théories sur la métrique de l'espace des couleurs. *Proc. 1st AIC congress "Color 69"*, Stockholm, 257-265
- Liebert, E. (1968), Farbabstand zwischen Istwert und Sollwert, seine Berechnung mit der MD Rechentafel. *Papier* 22, 376
- Luria, S. M., and Dimmick, F. L. (1964), Color discrimination. *Color Eng.* 2, 1: 14
- Luria, S. M., and Weissmann, S. (1965), Effect of stimulus duration on the perception of red-green and yellow-blue mixtures. *J. Opt. Soc. Amer.* 55, 1068-72
- MacAdam, D. L. (1942), Visual sensitivities to color differences in daylight. *J. Opt. Soc. Amer.* 32, 247-274
- MacAdam, D. L. (1963), Nonlinear relations of psychometric scale values to chromaticity differences. *J. Opt. Soc. Amer.* 53, 754-757
- MacAdam, D. L. (1964-'65), Analytical approximation for color metric coefficients.
 II. Friele approximation. *J. Opt. Soc. Amer.* 54, 249-256
 III. Optimization of parameters in Friele's formulas. *J. Opt. Soc. Amer.* 54, 1161-65
 IV. Smoothed modifications of Friele's formulas. *J. Opt. Soc. Amer.* 55, 91-95
- MacAdam, D. L. (1965), Specification of color differences. *Acta Chromatica* 1, 147-155
- MacAdam, D. L. (1965), Color measurement and tolerances. *J. Paint Technol.* 37, 1487-1531
- MacAdam, D. L. (1966), Smoothed versions of Friele's 1965 approximations for color metric coefficients. *J. Opt. Soc. Amer.* 56, 1784-1785
- MacAdam, D. L. (1969), Comment concerning scaling factors in color measurement formulas. *Color Eng.* 7, 6: 40
- MacAdam, D. L. (1970), *Sources of Color Science*. Cambridge, MIT Press
- MacAdam, D. L. (1969), Geodesic chromaticity diagram. *Proc. 1st AIC congress "Color 69"*, Stockholm, 293-301. *Farbe*, 18, 77-84
- MacAdam, D. L. (1971), Geodesic chromaticity diagram based on variances of color matching by 14 normal observers. *Appl. Optics* 10, 1-7
- MacAdam, D. L. (1971), Role of luminance increments in small color differences. This symposium
- McLaren, K. (1969), Scaling factors in color measurement formulas: A confusing situation. *Color Eng.* 7, 6: 38-44
- McLaren, K. (1969), The precision of textile colour matches in relation to colour difference measurements. *Proc. 1st AIC congress "Color 69"*, Stockholm, 688-708. *Farbe*, 18, 171-190
- McLaren, K. (1970), The Adams-Nickerson colour difference formula. (with conversion tables) *J. Soc. Dyers Col.* 86, 354
- McLaren, K. (1970), Colour passing—visual or instrumental. *J. Soc. Dyers Col.* 86, 389
- McLaren, K., and Coates, E. (1970), Recommended colour difference formula. *J. Soc. Dyers Col.* 86, 368
- Malkin, F. (1967), Colour tolerance studies with ceramic wall-tiles. *Proc. Symp. on Colour Measurement in Industry*, The Colour Group, London, 221

- Marsden, A. M. (1969), Brightness-a review of current knowledge. *Lighting Res. Technol.* 1, 171-181
- Marshall, W. J., and Tough, D. (1968), Colour measurement and colour tolerance in relation to automation and instrumentation in textile dyeing. *J. Soc. Dyers Col.* 84, 108
- Matveev, A. B. (1964), On the problems of uniform chromaticity color space construction. *Svetotechnika*, No. 12 (In Russian).
- Matveev, A. B., and Baljaeva, N. M. (1965), Uniform chromaticity color system, *Svetotechnika*. No. 1 (In Russian)
- Morton, T. H. (1969), ΔE by differential colorimetry in production dyeing. Proc. 1st AIC congress "Color 69", Stockholm, 709-716. *Farbe*, 18, 164-170
- Mudd, J. S. (1964), Correlation of colour tolerance determination with average observer assessments II. *Soc. Leather Trades Chem.* 48, 452
- Mudd, J. S., and Woods, M. (1970), Colour difference measurement. *J. Oil Col. Chem. Assn.* 53, 852
- Müller, U. (1969), Untersuchung der Helligkeitsempfindung unter Berücksichtigung der Änderung des Adaptationszustands des Auges. *Biophysik* 5, 276
- Nickerson, D. (1944), Summary of available information on small color difference formulas. In: Discussion Session on Small Color Differences, ISCC meeting 1944, reprinted by Am. Dyestuff Reprtr. 33, May, June, July
- Nickerson, D., and Stultz, K. F. (1944), Color tolerance specification. *J. Opt. Soc. Amer.* 34, 550-570
- Nimeroff, I. (1964), Spectral tristimulus values for the CIE (u , v , w) uniform spacing system. *J. Opt. Soc. Amer.* 54, 1365-1367
- Nimeroff, I. (1971), Does the 1964 CIE $U^*V^*W^*$ have a spectrum locus. This symposium
- Onley, J. W., and Klingberg, C. L. (1970), Saturation: peripheral or central? *J. Opt. Soc. Amer.* 60, 401-405
- Panek, D. W., and Stevens, S. S. (1966), Saturation of red: A prothetic continuum. *Percept. Psychophys.* 1, 59
- Parra, M. (1967), Quelques aspects nouveaux du seuil différentiel de couleur. *Couleurs* no. 65: 9-17
Couleurs no. 68: 11-15
- Parra, F. (1969) Seuils différentiels de couleur. Proc. 1st AIC congress "Color 69", Stockholm, 727-733
- Parra, F. (1971), Continuation of the study of colour thresholds.
- Pearson, D. E., Rubinstein, C. B., and Spivack, G. J. (1969), Comparison of perceived color in two-primary computer-generated artificial images with predictions based on the Helson-Judd formulation. *J. Opt. Soc. Amer.* 59, 644-658
- Pitt, I. T. (1967), The CIE colour difference formula applied to photographic colour reflection prints. Proc. Symp. on Colour Measurement in Industry, The Colour Group (Great Britain), 234
- Pokorny, J., and Smith, V. C. (1970), Wavelength discrimination in the presence of added chromatic fields. *J. Opt. Soc. Amer.* 60, 562-569
- Reuther, R. (1965), Messung des Simultankontrastes mit Hilfe des Gleichheitskriteriums, Proc. Int. Colour Meeting, Lucern, 1, 105-115
- Richards, W. (1966), Opponent-process solutions for uniform Munsell spacing. *J. Opt. Soc. Amer.* 56, 1110-1120
- Richter, K. (1969), Farbdifferenzformel des Gegenfarbensystems. *Farbe* 18, 207-220 (1969).
- Richter, K. (1969), New opponent colour concept for deriving Munsell hue and chroma as well as Evans' recent results. Proc. 1st AIC congress „Color 69", Stockholm, 403-417
- Richter, K. (1971), Description of colour attributes and colour differences. This symposium

- Richter, M. (1964), Eine Rechenscheibe zur Bestimmung der Dunkelstufe. *Farbe* 13, 157-161
- Richter, M. (1968), Another color difference formula. *Color Eng.* 6, 1: 38
- Richter, M., and Weise, H. (1969), Zum Vergleich verschiedener Farbabstandsformeln. Proc. 1st AIC congress "Color 69", Stockholm, 649-663. *Farbe*, 18, 149-163
- Robertson, A. R. (1969), A new determination of lines of constant hue. Proc. 1st AIC congress, "Color 69", Stockholm, 395-402
- Robinson, D. (1969), Acceptability of colour matches. *J. Oil Col. Chem. Ass.* 82, 15
- Ruddock, K. H. (1965), Integration processes in colour vision. Proc. 1st AIC congress, Lucern, 215-224. *Farbe*, 15 (1966), 63-72
- Saltzman, M. (1969), Comment concerning scaling factors in color measurement formulas. *Color Eng.* 7, 6: 40
- Sato, M., and Hioki, R. (1969), Modification of the (U^* , V^* , W^*) system, *J. Opt. Soc. Amer.* 59, 349-355
- Scheibner, H. (1964), Über Farben und ihre Metrik, *Lichttechnik* 16, 497
- Schultze, W. (1969), Umfassender Vergleich von sieben Farbabstandsformeln. Proc. 1st AIC congress "Color 69", Stockholm, 621-640. *Farbe*, 18, 105-130
- Schultze, W. (1971), Über die Brauchbarkeit von Farbabstandsformeln für die Festlegung von Farbtoleranzen (to be published). Also: This symposium
- Schultze, W., and L. Gall, Application of colour difference formulae to highly saturated colours differing only in lightness and saturation. (to be published)
- Sheppard, Jr. J. J. (1968), *Human Color Perception*, New York, Amer. Elsevier Publishers
- Shesternina, G. P., and Alyavdin, N. A. (1968), Employment of mathematical methods in processing determinations of small color differences of textile samples. *Industrial Laboratory* (translated from Russian) 34, 391
- Shirley, A. W. (1965), The systematic measurement of simultaneous contrast effects. Proc. Int. Colour Meeting, Lucern, 99-104. *Farbe*, 15 (1966), 57-62
- Shklover, D. A. (1957), The equicontrast colorimetric system. Symp. on Visual Problems of Colour, Teddington, p. 605-614; Also published by Chemical Publishing Co. New York, 1961 (there, see Vol. 2, p. 235)
- Siegel, M. H. (1964), Discrimination of color. IV. Sensitivity as a function of spectral wavelength, 410 through 500 $m\mu$. *J. Opt. Soc. Amer.* 54, 821-823
- Siegel, M. H. (1965), Color discrimination as a function of exposure time. *J. Opt. Soc. Amer.* 55, 566-568
- Siegel, M. H. (1969), The effect of forced choice on color discrimination. *Behav. Res. Meth. & Instru.* 1, 303
- Siegel, M. H., and Siegel, A. B. (1969), Colour discrimination and colour appearance. Proc. 1st AIC congress "Color 69", Stockholm, 391-394
- Siegel, M. H., and Siegel, A. B. (1971), Color name as a function of surround luminance and stimulus duration. *Percept. Psychophys.* 9, 140-144
- Siegel, M. H., and Siegel, A. B., (1971), A comparison of techniques for measuring hue appearance. *Behav. Res. Meth. & Instru.* 3, 1
- Simon, F. T., and Goodwin, W. J. (1958), Rapid graphical computation of small color differences. *Amer. Dyestuff Reprtr.* 47, 4: 105
- Simon, F. T. (1971), Industrial color tolerances by XI-ETA formula. This symposium
- Staes, K., and Verbrugghe, R. (1969), The role of the uniform colour space. Proc. 1st AIC congress "Color 69", Stockholm, pp. 1014-1022
- Steen, P. (1969), Experiments with estimation of perceptive qualitative color attributes. Proc. 1st AIC congress "Color 69", Stockholm, 369-376
- Steen, P., and Tonnquist, G. (1969), Colour difference ellipsoids in four systems. Proc. 1st AIC congress "Color 69", Stockholm, 717-726
- Stiles, W. S. (1946), A modified Helmholtz line element in brightness-colour space. *Proc. Phys. Soc. (London)* 58, 41-65
- Strocka, D. (1971), Color difference formulas and visual acceptability. *Appl. optics* 10, 1308-1313

- Suga, N. (1965), Evaluation of color difference between chromatic color chips with the gray scale for evaluating change in color. *Acta Chromatica 1*, 195
- Sugiyama, Y. (1964), Tables for rapid computations of small CIE color differences. *Acta Chromatica 1*, 118
- Sugiyama, Y. (1964), Note on the MacAdam nonlinear hypothesis, *J. Opt. Soc. Amer.* **54**, 127
- Sugiyama, Y. (1964), Review on measurements of and formulations of color difference. *Circ. Electrotechn. Lab. (Japan)* No. 156 (Japanese with English Abstract)
- Sugiyama, Y. (1964), Hyperbolic color-difference formulae. (Japanese with English Abstract) *Oyo Buturi 33*, 250
- Sugiyama, Y. (1965), Statistical tests for the triad-ratio comparisons on the color difference. *Proc. Int. Colour Meeting, Lucern*, 315-320
- Sugiyama, Y. (1967), A comparison of color-difference formulas. *Acta Chromatica 1*, 263
- Sugiyama, Y. (1968), Transformations of color discrimination ellipses. *J. Illum. Eng. Inst. (Japan)* **52**, 12
- Sugiyama, Y. (1969), Note on the Brown's paper. *Acta Chromatica 2*, 48
- Sugiyama, Y. (1969), Introduction of color science, 2. Color solids and color difference formulas, *Color Space No. 52/53*, 29. 3. Threshold, color-difference and its application and others. *Color space No. 54*, 29
- Sugiyama, Y. (1970), Bibliography on uniform color scales. *Color Space No. 59*, 10
- Sugiyama, Y. (1970), Introduction of color science: Supplement, new Euclidean geometry and color space. *Color Space No. 56/57*, 29
- Sugiyama, Y. (1971), Comparison of color-difference equations by using the ratio comparison. To be published
- Sugiyama, Y. and Fuwa, M. (1966), 1960 CIE-UCS chromaticities of the Munsell renotations and their diagrams. *Bull. Electrotechnical Lab. (Japan)* **30**, 296 (Japanese with English abstract)
- Sugiyama, Y., and Wright, H. (1963), Paired comparisons of color differences. *J. Opt. Soc. Amer.* **53**, 1214-1222
- Sugiyama, Y., and Wright, H. (1964), Ratio comparisons of color differences. *J. Opt. Soc. Amer.* **54**, 75-78
- Taguti, R., and Sato, M. (1962), Exponential color coordinate system. *Acta Chromatica 1*, 19
- Takasaki, H. (1966), Lightness change of grays induced by change in reflectance of gray background. *J. Opt. Soc. Amer.* **56**, 504-509
- Terstiege, H. (1971), The influence of the fundamental primaries on chromatic adaptation and color-difference evaluation under different illuminants. This symposium
- Thielert, R., and Schliemann, G. (1964), Die subjektive Beurteilung der Leuchtkraft von Tagesleuchtfarben und ihre Beziehung zu farbmetrischen Grössen. *Proc. 1st AIC congress "Color 69"*, Stockholm, 427-436
- Thurner, K., and Walther, V. (1969), Untersuchungen zur Korrelation von Farbabstands-bewertungen auf visuellem Wege und über Farbdifferenzformeln. *Proc. 1st AIC congress "Color 69"*, Stockholm, 671-687. *Farbe, 18*, 191-206
- Tonnquist, G. (1965), A comparison between symmetrical and equi-spaced hue circles. *Proc. Int. Colour Meeting, Lucern*, 376-388. *Farbe, 15* (1966) 311-323
- Trabka, E. A. (1968), On Stile's line element in brightness-color space and the color power of the blue. *Vision Res.* **8**, 113-134
- Tsukada, I., and Minato, S. (1965), Perceptually uniform steps in the hue circle (A new color circle having perceptually uniform steps). (A Japanese Journal, unidentified)
- Tsukada, I., Minato, S., and Miyazaki, Y. (1966), Interpolation of the chromaticity coordinates (x , y) of the Munsell colors at 100 levels of Munsell Value
Part I (1966)
Part II (1966)
(A Japanese Journal, unidentified)
- Valberg, A., and Holtsmark, T. (1971), Similarity between JND-curves for complimentary optimal colours. This symposium

- Vos, J. J., and Walraven, P. L. (1972), An analytical description of the line element in the zone-fluctuation model of colour vision. I, II. *12 Vision Res.* 12, 1327-44; 1345-65
- Vos, J. J., and Walraven, P. L. (1969), A zone-fluctuation line element describing colour discrimination. *Proc. 1st AIC congress "Color 69"*, Stockholm, 291-292
- Vos, J. J., and Walraven, P. L. (1971), A zone-fluctuation line element describing colour discrimination. *This symposium*
- Walraven, P. L. (1965), A zone theory of colour vision. *Proc. Int. Colour Meeting, Lucern*, 137-140. *Farbe*, 15 (1966), 17-20
- Walraven, P. L. (1967), A uniform chromaticity diagram based upon a square root transformation of the colour space. *Proc. C.I.E. 16th Session, Washington 1967*, C.I.E. Publ. No. 14A (1968) Vol. A, 106-111
- Walraven, P. L., and Bouman, M. A. (1966), Fluctuation theory of colour discrimination of normal trichromats. *Vision Res.* 6, 567-586
- Warren, R. M. (1967), Quantitative judgments of color: the square root rule. *Percept. Psychophys.* 2, 448-452
- Warren, R. M., and Poulton, E. C. (1966), Lightness of grays: Effects of background reflectance. *Percept. Psychophys.* 1, 145
- Warren, R. M., and Warren, R. P. (1965), Lightness, brightness, and distance. *Proc. Int. Colour Meeting, Lucern*, 351-356. *Farbe*, 15 (1966) 271-276
- Weise, H. (1965), Vergleich dreier skandinavischer Farbsysteme mit dem der DIN-Farbenkarte. *Proc. Int. Colour Meeting, Lucern*, 389-398. *Farbe*, 15 (1966), 277-286
- Weissmann, S., and Kinney, J. A. S. (1965), Relative yellow-blue sensitivity as a function of retinal position and luminance level. *J. Opt. Soc. Amer.* 55, 74-77
- Wright, H. (1965), Precision of color differences derived from a multidimensional scaling experiment. *J. Opt. Soc. Amer.* 55, 1650-1655
- Wright, W. D. (1941), The sensitivity of the eye to small colour differences. *Proc. Phys. Soc. (London)* 53, 93-112
- Wright, W. D., (1964), *The Measurement of Colour*. 4th ed., London, Hilger
- Wyszecki, G. (1963), Proposal for a new color-difference formula. *J. Opt. Soc. Amer.* 53, 1318-1319
- Wyszecki, G. (1965), Matching color differences. *J. Opt. Soc. Amer.* 55, 1319-1324
- Wyszecki, G. (1965), The measurement of color differences. *Proc. Int. Colour Meeting, Lucern*, 287-301. *Farbe*, 14, 67-79
- Wyszecki, G. (1967), Correlate for lightness in terms of CIE chromaticity coordinates and luminous reflectance. *J. Opt. Soc. Amer.* 57, 254-257
- Wyszecki, G. (1968), Recent agreements reached by the Colorimetry Committee of the Commission Internationale de l'Eclairage. *J. Opt. Soc. Amer.* 58, 290-292
- Wyszecki, G. and Fielder, G. H. (1971), New color-matching ellipses. *J. Opt. Soc. Amer.* 61, 1135-1152 Also *This symposium*
- Wyszecki, G., and Fielder, G. H. (1971), Color-difference matches. *J. Opt. Soc. Amer.* 61, 1501-1513
- Wyszecki, G., and Stiles, W. S. (1967), *Color Science. Concepts and methods, quantitative data and formulas*. New York etc., Wiley
- Wyszecki, G., and Wright, H. (1965), Field trial of the 1964 CIE color-difference formula. *J. Opt. Soc. Amer.* 55, 1166-1174
- Yonemura, G. T. (1970), Opponent-color-theory treatment of the CIE 1960 (u , v) diagram: Chromaticness difference and constant-hue loci. *J. Opt. Soc. Amer.* 60, 1407-1409
- Yonemura, G. T., and Kasuya, M. (1969), Color discrimination under reduced angular subtense and luminance, *J. Opt. Soc. Amer.* 59, 131-135

31 A SURVEY OF SOME CURRENT COLOUR DIFFERENCE FORMULAE

L. F. C. FRIELE

*Fibre Research Institute TNO
Delft, Netherlands*

A. Color difference formulas based on projective transformations of the CIE chromaticity diagram combined with a lightness scale

NBS color difference formula (1942) J. Opt. Soc. Amer. 32, 509-538

$$(\Delta E)^2 = (700)^2 Y_m^{\frac{1}{2}} [(\Delta \alpha)^2 + (\Delta \beta)^2] + [100 \Delta(Y^{\frac{1}{2}})]^2 \text{ NBS units or judds.}$$

$$0 < Y \leq 1 \quad Y_m = \frac{Y_1 + Y_2}{2}$$

$$\alpha = \frac{2.4266 x - 1.3631 y - 0.3214}{x + 2.2633 y + 1.1054}$$

$$\beta = \frac{0.5710 x + 1.2447 y - 0.5708}{x + 2.2633 y + 1.1054}$$

For glossy samples and for conditions where the samples to be compared are separated, modifications of the formula are indicated.

This formulae is indicated as "JH" by Coates *et al*, Schultze, and Jaeckel.

Hunter *L a b system* (1958) J. Opt. Soc. Amer. 48, 985-995

Based on Adams chromatic valence system (J. Opt. Soc. Am. 32 (1942) 168)

$$\Delta E^2 = (\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2$$

$$L = 10 Y^{\frac{1}{3}} \quad (0 < Y \leq 100)$$

$$a = 1.75 L \frac{(1.02 X - Y)}{Y}$$

$$b = 0.7 L \frac{(Y - 0.847 Z)}{Y}$$

CIE color difference formula (1964) J. Opt. Soc. Amer. 53, (1963) 1318-1319

$$(\Delta E)^2 = (\Delta W^*)^2 + (\Delta U^*)^2 + (\Delta V^*)^2$$

$$W^* = 25 Y^{\frac{1}{3}} - 17 \quad 1 \leq Y \leq 100$$

$$U^* = 13 W^* (u - u_0)$$

$$V^* = 13 W^* (v - v_0)$$

$$u = \frac{4x}{-2x + 12y + 3} \quad v = \frac{6y}{-2x + 12y + 3}$$

u_0 and v_0 are the coordinates of the illuminant.

This formulae is indicated as "C 64" by Jaekel.

B. Color difference formulas based on the Munsell Color System

The Munsell system specifies a color in terms of hue (H) value (V) and chroma (C). For the interrelation of Munsell specifications and CIE specifications, a smoothed graphical system was developed, defining the Munsell renotation system [J. Opt. Soc. Amer. 33, (1943) 385-418].

Nickerson (1936) index of fading. Textile Res. 6 509-514

$$\Delta E = \frac{C}{5} 2 \Delta H + 6 \Delta V + 3 \Delta C$$

Balinkin (1939) formula. Amer. J. Psychol. 52, 428-448

$$(\Delta E)^2 = \left(\frac{2C}{5} \Delta H \right)^2 + (6 \Delta V)^2 + \left(\frac{20}{\pi} \Delta C \right)^2$$

Godlove (1951) formula. J. Opt. Soc. Amer. 41, 760-772

$$(\Delta E)^2 = 2 C_1 C_2 [1 - \cos 3.6^\circ \Delta H] + (\Delta C)^2 + (4 \Delta V)^2$$

Munsell Renotation formula (1967). Color Eng. 8: 2 (1970) 36-52

$$\frac{(\Delta E)^2}{f_s^2} = 2 f_h C_1 C_2 (1 - \cos 3.6^\circ \Delta H) + (\Delta C)^2 + (4 \Delta V)^2$$

$$f_h = \left[\frac{4}{3 - \cos 3.6^\circ \Delta H} \right]^2$$

$$f_s = \frac{15 + [\bar{C}^2 + 16 (\bar{V} - V_s)^2]^{\frac{1}{2}}}{.5 + [\bar{C}^2 + 16 (\bar{V} - V_s)^2]^{\frac{1}{2}}}$$

$$\bar{V} = \frac{V_1 + V_2}{2} \quad \bar{C} = \frac{C_1 + C_2}{2}$$

V_s = Value of the surround

Adams-Nickerson-Stultz (1944) chromatic valence formula.

J. Opt. Soc. Amer. 34, 550-570

Based on Adams chromatic valence system

$$(\Delta E)^2 = (0.5 \Delta V)^2 + \left[\Delta \left\{ V \left(\frac{1.02 X}{Y} - 1 \right) \right\} \right]^2 + \left[0.4 \Delta \left\{ V \left(\frac{0.847 Z}{Y} - 1 \right) \right\} \right]^2$$

Adams-Nickerson-Stultz (1944) chromatic value formula.

J. Opt. Soc. Amer. 34, 550-570

Based on Adams chromatic value system (J. Opt. Soc. Am. 32, (1942) 168)

$$\frac{(\Delta E)^2}{K^2} = (0.23 \Delta V_y)^2 + [\Delta(V_x - V_y)]^2 + [0.4 \Delta(V_z - V_y)]^2$$

V_x , V_y and V_z result by applying the Munsell value function to 1.02 X , Y and 0.847 Z respectively.

Tables for the interrelation of (X, Y, Z) and (V_x, V_y, V_z) see:

Judd-Wyszecki: Color in business science and industry, 2nd Ed. 1963 Wiley N.Y.

Wyszecki-Stiles: Color Science 1967 Wiley N.Y.

McLaren : J. Soc. Dyers. and Col. 86 (1970) 356-366

The constant K defines the ΔE unit. There is some discussion on the value of K : See K. McLaren, Color Eng. 7: 6 (1969) 38-44

This formulae is indicated as "Adams" by Malkin and Dinsdale, as "ANS" by Schultze, as "ACV" by Jaeckel and Coates *et al*, and as "ANLAB" by McLaren.

Saunderson-Milner (1946). J. Opt. Soc. Amer. 36, 36-42

$$(\Delta E)^2 = (K \Delta V)^2 + [\Delta(9.37 + 0.79 \cos \Theta)(V_x - V_y)]^2 + [0.4 \Delta(8.33 + 2.18 \sin \Theta)(V_z - V_y)]^2$$

$$\tan \Theta = \frac{0.4(V_z - V_y)}{V_x - V_y}$$

Cube-root formula. (1958). J. Opt. Soc. Amer. 48, 736-740

$$(\Delta E)^2 = [25.29 \Delta(G^\dagger)]^2 + [106 \Delta(R^\dagger - G^\dagger)]^2 + [42.34 \Delta(G^\dagger - B^\dagger)]^2$$

$$L- \quad R = 1.1084 X + 0.0852 Y \quad \Delta \quad 0.1454 Z$$

$$G = -0.0010 X + 1.0005 Y + 0.0004 Z$$

$$B = -0.0062 X + 0.0394 Y + 0.8192 Z$$

The RGB transformation was suggested by Reilly at the CIE Meeting 1963.

This formulae is indicated as $\sqrt[3]{\quad}$ by Malkin and Dinsdale.

Friele (1969). Proc. 1st AIC congress "Color '69" Stockholm, 275-290

$$\begin{aligned}
 (\Delta E)^2 &= [K \Delta(Y^{\frac{1}{3}})]^2 + [\Delta C_{r-g} - 0.5 \Delta C_{y-b}]^2 + [\Delta C_{y-b}]^2 \\
 C_{y-b} &= \frac{Y - S_n}{0.155 Y^{2/3}} \quad \text{for } S_n > S_1 \\
 C_{y-b} &= \frac{Y - S_1}{0.155 Y^{2/3}} + \frac{\ln S_1/S_n}{0.5} \quad \text{for } S_n < S_1 \left\} S_1 = \frac{0.155 Y^{2/3}}{0.5} \\
 C_{r-g} &= 65 (P_n^{\frac{1}{3}} - P_o^{\frac{1}{3}}) \quad \text{for } P_n < Q_n \\
 C_{r-g} &= -105 (Q_n^{\frac{1}{3}} - Q_o^{\frac{1}{3}}) \quad \text{for } P_n > Q_n \\
 P_o = Q_o &= 0.957 Y + 0.044 S_n \\
 P_n &= 0.742 X_c + 0.391 Y_c - 0.100 Z_c \\
 Q_n &= -0.457 X_c + 1.305 Y_c + 0.122 Z_c \\
 S_n &= 0.847 Z_c \\
 0 < Y &\leq 1
 \end{aligned}$$

C. Color difference formulas, based on threshold data

Simon-Goodwin (1958) Amer. Dyestuff. Rep. 47, 105-112

and Foster (1966) Color Eng. 4:1, 17-19

Graphical computation systems

Both are based on the MacAdam threshold ellipses (J. Opt. Soc. Amer. 32 (1942), 247-274) and the interpolated ellips coefficients (J. Opt. Soc. Amer. 33 (1943), 18-26).

The Simon-Goodwin and Foster systems are obtainable from Diano Corp. P.O. Box 231, Foxboro, Mass. 02035.

1965 Friele formula J. Opt. Soc. Amer. 55, 1314-1319

$$\begin{aligned}
 (\Delta E)^2 &= \left[\frac{l \Delta L}{a} \right]^2 + \left[\frac{\Delta C_{r-g}}{a} - f \frac{\Delta C_{y-b}}{b} \right]^2 + \left[\frac{\Delta C_{y-b}}{b} \right]^2 \\
 \Delta L &= RG (R^2 + G^2)^{-\frac{1}{2}} \left[\frac{\Delta R}{G} + \frac{\Delta G}{R} \right] \\
 \Delta C_{y-b} &= RG (R^2 + G^2)^{-\frac{1}{2}} \left[\frac{\Delta R}{R} - \frac{\Delta G}{G} \right] \\
 \Delta C_{y-b} &= RGB (R^2 + G^2)^{-1} \left[\frac{\Delta R}{G} + \frac{\Delta G}{R} \right] - \Delta B
 \end{aligned}$$

$$R = 0.679 X_c + 0.491 Y_c - 0.133 Z_c$$

$$G = -0.495 X_c + 1.371 Y_c + 0.097 Z_c$$

$$B = 0.847 Z_c$$

$$\text{High luminance: } Y > 0.5 ft-L \quad B > 1.0$$

$$a = \alpha R \quad \text{for } R > G; \quad a = \alpha G \quad \text{for } G > R$$

$$b = \beta [B^2 + (pY)^2]^{\frac{1}{2}}$$

$$\alpha = 0.0035 \quad \beta = 0.0175 \quad p = 0.4$$

$$\text{Threshold differences: } l = 0.25 \quad f = 0$$

$$\text{Commercial tolerances: } l = 0.1 \quad f = 0.75$$

$$\text{Low luminance: } Y < 0.25 ft-L \quad B < 1.0$$

$$a = \alpha R^{\frac{1}{2}} \quad \text{for } R > G; \quad a = \alpha G \quad \text{for } G > R$$

$$b = \beta [B + pY]^{\frac{1}{2}}$$

$$\alpha = 0.0035 \quad \beta = 0.03 \quad p = 0.4$$

$$\text{Threshold differences: } l = 0.35 \quad f = 0$$

FMC-1 (1967) J. Opt. Soc. Amer. 57, 537-541

1965 Friele formula, modified by MacAdam, optimized by Chickering. The optimization is valid for the MacAdam (1942) ellipses.

$$(\Delta E)^2 = \left[\frac{l \Delta^{\frac{1}{2}}}{a} \right]^2 + \left[\frac{\Delta C_{r-g}}{a} - f \frac{\Delta C_{y-b}}{b} \right]^2 + \left[\frac{\Delta C_{y-b}}{b} \right]^2$$

$$\Delta L = PQ (P^2 + Q^2)^{-\frac{1}{2}} \left(\frac{\Delta P}{Q} + \frac{\Delta Q}{P} \right)$$

$$\Delta C_{r-g} = PQ (P^2 + Q^2)^{-\frac{1}{2}} \left(\frac{\Delta P}{P} - \frac{\Delta Q}{Q} \right)$$

$$\Delta C_{y-b} = PQS (P^2 + Q^2)^{-1} \left(\frac{\Delta P}{Q} + \frac{\Delta Q}{P} \right) - \Delta S$$

$$P = 0.724 X + 0.382 Y - 0.098 Z$$

$$Q = -0.490 X + 1.370 Y + 0.1276 Z$$

$$S = 0.686 Z$$

$$a^2 = \alpha^2 (P^2 + Q^2) \left[1 + N \frac{P^2 Q^2}{P^4 + Q^4} \right]^{-1}$$

$$b^2 = \beta^2 [S^2 + (pY)^2]$$

$$\alpha = 0.00416 \quad \beta = 0.0176 \quad p = 0.4489 \quad f = 0$$

$$N = 2.73 \quad l = 0.279$$

This formulae is indicated as "FMC" by Jaeckel.

FMC-2 (1967) J. Opt. Soc. Amer. 61 (1971), 118-122

$$(\Delta E)^2 = \left(K_2 \frac{l \Delta L}{a} \right)^2 + K_1^2 \left[\frac{\Delta C_{r-g}}{a} - f \frac{\Delta C_{y-b}}{b} \right]^2 + K_1^2 \left[\frac{\Delta C_{y-b}}{b} \right]^2$$

$$K_1 = 0.55669 + 0.049434 Y - 0.82575 \cdot 10^{-3} Y^2 + 0.79172 \cdot 10^{-5} Y^3 \\ - 0.30087 \cdot 10^{-7} Y^4$$

$$K_2 = 0.17548 + 0.027556 Y - 0.57262 \cdot 10^{-3} Y^2 + 0.63893 \cdot 10^{-5} Y^3 \\ - 0.26731 \cdot 10^{-7} Y^4$$

$$0 < Y < 100$$

Otherwise as for FMC-1

This formulae is indicated as "MFC-2" by "Coates *et al*" and as "Fkk" by Jaeckel.

MacAdam ($\xi\eta$) geodesic formula (1969) Farbe 18 (1969) 77-84

Appl. Optics 10 (1971) 1-7

$$(\Delta E)^2 = (K_1 \Delta C)^2 + \left(K_2 \frac{\Delta Y}{Y} \right)^2$$

$$(\Delta C)^2 = (\Delta \xi)^2 + (\Delta \eta)^2$$

$$\xi = \xi(a, b) \quad \eta = \eta(a, b)$$

a and b are coordinates, resulting from projective transformations of the CIE chromaticity chart.

D. Other Color difference formula

DIN 1955 J. Opt. Soc. Amer. 45, 223-226

Color Eng. 6 (1968): 1, 38-40

$$(\Delta E)^2 = \left(\frac{10-D}{9} \cdot \frac{S}{6} \Delta T \right)^2 + \left(\frac{10-D}{9} \Delta S \right)^2 + (\Delta D)^2$$

$$D = 10 - 6.1723 \log \left(40.7 \frac{Y}{Y_0} + 1 \right) \text{ (Darkness degree)}$$

Y_0 is the maximum luminous reflectance for the chromaticity in question (optimal color).

T and S are correlates for hue and saturation resp. and can be read from graphs.

Author Index

Numbers refer to *articles*, not to pages. Bold-face numbers relate to the articles themselves, light-face numbers to references.

- Abramov, I., 6
Abney, W. de W., 1
Adam, M., 30
Adams, E. Q., 13, 23, 31
Adrian, W., 30
Alexander, J. F., 26
Alexander, J. V., 1, 6, 25
Allen, E., 30
Alpern, M., 1
Alyavdin, N. A., 30
American Society of Testing Materials, 23
Ampt, C. G. F., 10, 29
Atherton, E., 23, 24
Balinkin, I. A., 21, 22, 30, 31
Baljaeva, N. M., 30
Barlow, H. B., 6, 29
Bartleson, C. J., 18, 26, 30
Beare, A. C., 30
Beasley, J. K., 12
Bedford, R. E., 4
Bell, J. R., 30
Bellamy, B. R., 11
Benson, K. B., 18
Berger, A., 23, 30
Best, R. P., 30
Bilger, H., 5
Billmeyer, Jr., F. W., 12, **19**, 20, 27, 30
Blackwood, N. K., 30
Blottiau, F., 30
Bodmann, H. W., 30
Bond, M. E., 20
Bongard, M. M., 3
Bouma, H., 20, 30
Bouma, P. J., 1, 5, 9
Bouman, M. A., 1, 4, 6, 13, 29, 30
Boynton, R. M., 1, **2**, 7, 9, 29, 30
Breneman, E. J., 18, 30
Brewer, W. L., 9, 18
Brickwedde, F. G., 13, 16
Bridgeman, T., 30
British Standards 950, prt 1, 21
Brookes, A., 23, 30
Brocklebank, R. W., 13
Brodhun, E., 1
Brown, W. R. J., 1, 2, 6, 14, 15, 21, 24, 27, 30
Burch, J. M., 3
Burnham, R. W., 2, 26, 30
Campbell, E. D., **19**, 30
Chapanis, A., 18
Checchi, R., 30
Chickering, K. D., 6, 13, 14, 15, 19, 23, 26, 27, 30, 31
Chumbly, J. I., 1, 6
CIE, 9, 20, 30
Clarke, F. J. J., 1, 3, 30
Coates, E., 22, 23, **24**, 30
Colour Measurement Committee, 25
Connors, M. M., 30
Cornsweet, T. N., 13, 30
Crawford, B. H., 3, **8**, 16, 21, 30
Curtis, D. W., 30
Davidson, H. R., 21, 22, 23, 24, 25, 27, 30
Davidson, S. L., 30
Day, S., 22, 23, **24**, 30
De Lange Dzn., H., 4
DeValois, R. L., 6
De Vries, H., 1, 9
Dieterici, C., 1, 13
Dimmick, F. L., 30
Dinsdale, A., **21**, 23, 30
Dobelle, W. H., 6
Donley, N. J., 1, 7
Dresler, A., 4
Ekman, G., 11
Evans, R. M., 12, 30
Farnsworth, D., 13, 30
Federov, N. T., 9
Fielder, G. H., **15**, 30
Fiorenzi, G., 30
Fisher, R. A., 23
Flock, H. R., 30
Foss, C. E., 10
Foster, R. S., 30, 31
Friede, E., 21, 22, 23, 24, 25, 27, 30
Friele, L. F. C., 1, 13, 14, 19, 20, 26, **27**, 30, **31**
Fry, G. A., 3, 13
Fujii, K., 30
Fukuda, T., 30
Fuwa, M., 30
Gall, L., 14, 22, 27, 30
Ganz, E., 30
Garrett, D. A., 23, 24
Gasser, M., 5
Gaylord, H. A., 30
Gibson, N., 30
Gibson, K. S., 1, 10, 13
Gillman, C. B., 1, 6
Glasser, L. G., 1, 13, 21, 23, 26, 31
Glenn, J. J., 13
Godlove, I. H., 10, 24, 31
Goodwin, W. J., 21, 23, 26, 30, 31
Gordon, J., 30
Graham, C. H., 30

- Granville, W. C., 10
 Guilford, J. P., 24
 Guth, S. L., 1, 6, 7
 Guttman, L., 11
 Halsey, R. M., 18
 Hanes, R. M., 26
 Hanlon, J. J., 23, 30
 Hård, A., 12, 30
 Hardy, A. C., 13
 Harvey, Jr., L. O., 30
 Hecht, S., 1
 Helm, C. E., 30
 Helmholtz, H. von, 1
 Hemmendinger, H., 30
 Herzberg, P. A., 2
 Hioki, R., 30
 Hofmann, K. D., 5
 Holtsmark, T., 5, 30
 Hotelling, H., 23
 Hubel, D., 6
 Huff, J. E., 7
 Hunt, R. W. G., 18, 30
 Hunter, R. S., 1, 11, 20, 23, 26, 31
 Hurvich, L. M., 6, 12, 30
 Hutchings, D. M., 23
 Hyman, R., 11
 Ikeda, M., 1
 Illing, A. M., 30
 Indow, T., 2, 11, 30
 Ishak, I. G. H., 20, 30
 Jacobs, G. H., 30
 Jacquemart, J., 30
 Jaeckel, 23, 25, 30
 Jäkel-Hartenstein, B., 30
 Jameson, D., 6, 12, 30
 Johnston, R., 20, 30
 Jonckheere, J., 30
 Judd, D. B., 1, 4, 5, 6, 7, 8, 10, 13, 17, 18,
 20, 24, 26, 30, 31
 Justova, E. N., 9
 Kaiser, P. K., 2, 7
 Kambe, N., 4, 30
 Kanazawa, K., 11, 30
 Kandel, G. L., 19, 30
 Kaneko, T., 30
 Kawai, T., 11
 Kelly, K. L., 13, 17
 Killian, J. T., 13
 Kinney, J. A. S., 30
 Klingberg, C. L., 30
 König, A., 1, 13
 Kornerup, A., 30
 Kowaliski, P., 12, 18, 30
 Krantz, D. H., 30
 Kruskal, J. B., 2, 11
 Kuehni, R., 21, 23, 27, 28, 31
 Kustarev, A. K., 9
 Ladd, J. H., 17, 18
 Ladd-Franklin, C., 6
 Land, E. H., 21
 Leebeek, H. J., 4
 LeGrand, Y., 1, 9, 30
 Liebert, E., 30
 Lingoes, J. C., 11
 Little, A. C., 23
 Lobanova, P. A., 9
 Lodge, H. R., 7
 Luce, R. D., 10
 Luria, S. M., 30
 Luther, R., 5
 MacAdam, D. L., 1, 2, 6, 9, 13, 14, 15, 19,
 20, 21, 23, 24, 26, 27, 30, 31
 MacMillan, J., 19, 30
 MacNichol, Jr., E. F., 6
 Malkin, F., 21, 23, 30
 Marks, W. B., 6
 Marrocco, R. T., 1, 7
 Marsden, A. M., 30
 Marshall, W. J., 23, 24, 30
 Matsushima, L., 11, 30
 Matveev, A. B., 30
 McCann, J. J., 21
 McCree, K. J., 4, 6
 McKinney, A. H., 13, 21, 23, 26, 31
 McLaren, K., 14, 20, 21, 22, 23, 24, 25,
 27, 30, 31
 McNicholas, H. J., 10
 Miescher, K., 5
 Milner, B. J., 1, 23, 31
 Minato, S., 30
 Miyazaki, Y., 30
 Moon, P., 13
 Morton, T. H., 23, 30
 Mudd, J. S., 22, 23, 30
 Müller, U., 30
 Munsell, A. E. O., 10
 Muth, E. J., 13, 14
 Neun, M. E., 30
 Newhall, S. M., 1, 10, 11, 13, 24, 31
 Nickerson, D., 1, 10, 11, 13, 20, 22, 23, 24,
 30, 31
 Nimeroff, I., 17, 30
 Nyberg, N. D., 9
 Ohsumi, K., 2, 11
 Onley, J. W., 2, 30
 Panck, D. W., 30
 Parra, F., 16, 30
 Patterson, M. M., 1, 6
 Pearson, D. E., 30
 Penciolelli, G., 30
 Persels, C. G., 13, 14
 Pinney, J. E., 17, 18
 Pitt, F. H. G., 6, 9
 Pitt, I. T., 30
 Poincaré, H., 10
 Pokorný, J., 30

- Poulton, E. C., 30
 Priest, I. G., 10, 13, 16
 Provost, 23, **24**, 30
 Ramsay, J. O., 2, 11
 Rautian, C. N., 9
 Reilly, C. D., 13, 21, 23, 26, 31
 Reuther, R., 30
 Richards, W., 30
 Richter, K., 5, **12**, 13, 30
 Richter, M., 20, 30, 31
 Riemann, B., 10
 Rigg, B., 22, 23, **24**, 30
 Robertson, A. R., 13, 20, 30
 Robinson, F. D., 22, 30
 Rose, A., 1
 Roylance, S., **20**
 Rubenstein, C. B., 30
 Ruddock, K. H., 30
 Rushton, W. A. H., 1, 6
 Saltzman, M., 30
 Sanders, C. L., 18
 Sato, M., 30
 Saunderson, J. L., 1, 23, 31
 Savoie, R. E., 13
 Schafer, W., 30
 Scheibner, H., 30
 Schliemann, G., 30
 Schnelle, P. D., 13, 21, 23, 26, 31
 Schrödinger, E., 1, 6, 13, 14
 Schultze, W., 14, 17, **22**, 23, 25, 27, 30
 Scofield, F., 1, 23
 Sheldon, J. A., 12
 Shepard, R. N., 2, 11
 Sheppard, Jr., J. J., 30
 Shesternina, G. P., 30
 Shiose, T., 11
 Shirley, A. W., 30
 Shklover, D. A., 1, 30
 Siegel, A. B., 30
 Siegel, M. H., 30
 Silberstein, L., 1
 Simon, F. T., 21, 23, **26**, 30, 31
 Sloan, L. L., 10
 Smith, R., 30
 Smith, V. C., 30
 Spencer, D. E., 13
 Speranskaya, N. I., 3
 Sperling, H. G., 9
 Spivack, G. J., 30
 Staes, K., 30
 Steen, P., 30
 Steffen, D., 9
 Stevens, S. S., 11, 30
 Stiles, W. S., **1**, 2, 3, 6, 9, 15, 30, 31
 Strocka, D., 14, 22, 23, 30
 Stultz, K. F., 20, 22, 23, 24, 30, 31
 Suga, N., 30
 Sugiyama, Y., 30
 Swenholt, B. K., 12, 30
 Szekeres, G., 9
 Taguti, R., 30
 Takagi, C., 2, 30
 Takasaki, H., 30
 Teller, D. Y., 30
 Terstiege, H., **9**, 30
 Thielert, R., 30
 Thomson, L. C., 9
 Thurner, K., 20, 21, 22, 24, 25, 27, 30
 Tilleard, D. L., 20
 Tonnquist, G., 30
 Torgerson, W. S., 11
 Torii, S., 1
 Tough, D., 23, 24, 30
 Trabka, E. A., 1, 29, 30
 Trezona, P. W., **3**
 Troy, D. J., 21
 Tsukada, I., 30
 Tyler, J. E., 13
 Tyndall, E. P. T., 1
 Uchizono, T., 11
 USA Standards, 18
 Valberg, A., **5**, 12, 30
 Van Brussel, H. J. J., 20, 30
 Van der Horst, G. J. C., 4
 Van der Velden, H. A., 1
 Van Meeteren, A., 6, 29
 Verbrugghe, R., 30
 Vickerstaff, T., 23, 24
 Von Kries, J., 9
 Vos, J. J., 1, **6**, 27, 29, 30
 Wagner, H. G., **2**, 7, 30
 Walraven, P. L., 1, 4, **6**, 13, 27, 29, 30
 Walters, H. V., 1, 9
 Walther, V., 20, 21, 22, 24, 25, 27, 30
 Warburton, F. L., 22, 24
 Ward, C. D., 23
 Warren, R. M., 30
 Warren, R. P., 30
 Weale, R. A., 7
 Weise, H., 30
 Weisenhorn, P., 5
 Weissman, S., 30
 Well, A., 11
 Werner, A. J., 17
 Whitten, D. N., 29
 Wiesel, T. N., 6
 Williams, E. J., 23
 Willmer, E. N., 4
 Wilson, M. H., 13
 Witzel, R. F., 2, 30
 Woods, M., 22, 23, 30
 Wright, H., 30
 Wright, W. D., 1, 4, 6, 9, 13, 17, 20, 30
 Wyszecski, G., 1, 2, 4, 6, 7, 13, 14, **15**, 17,
 18, 19, 20, 21, 23, 24, 26, **30**, 31
 Yonemura, G., 4, 13, 30