
Theory and practice of colour vision testing : A review. Part 2

R. LAKOWSKI

Visual Laboratory, Department of Psychology, Edinburgh University

Lakowski, R. (1969). *Brit. J. industr. Med.*, **26**, 265-288. **Theory and practice of colour vision testing: A review. Part 2.** The more modern tests of colour confusion, hue discrimination, metamerism, and special aptitudes measure a wide variety of attributes. Norms, including variations among normal as well as among defective observers, are given for each test, along with any other information we require to understand them. Finally there is the problem of choice, and factors influencing it, of which tests should be used in different circumstances. Some advice is given on the most useful combinations of tests.

Colour confusion – the pseudo-isochromatic tests (PIC)

Farnsworth (1955) pointed out that complete specification of stimuli in terms of the physical standards and psychological variables affecting vision and a better definition of experimental populations are needed if we are to study defective and normal vision adequately. This advice applies particularly to pseudo-isochromatic tests. Here the problem is quite complex for theoretically these tests should only be used to divide people into normals and defectives, but, because they are thought to be simple and are easily administered, they are in fact used as a kind of universal colour vision test from which more detailed information even about discrimination is extracted. But these tests are the least understood. In design they are very complex though, it must be admitted, ingenious, especially when we consider that many were brought out before accurate colorimetric measurements were available. Yet when examined in detail these plates are often far from ideally designed, and it is because of these faults that we can extract this additional information.

Then there is the problem of scoring; of what criteria should be used to classify defects. This has not yet been solved in spite of the numerous papers on this subject, some of them very involved indeed. At present, therefore, it is always better to look

upon information from PIC tests as providing a *probable* but not a certain diagnosis.

These tests are called pseudo-isochromatic because normal observers can distinguish between individual colours in the plate and read the intended figures, and therefore they are isochromatic only to defectives (see section on Colour measurement). Basically each of them consists of a number of plates where surface colours are employed in the form of dots to show a multi-coloured figure against a multi-coloured background. Variations in the size of dot (from 2 to 10 mm.), and the employment of arabic numerals and geometric designs for the figure, introduce problems of subtense and form-perception. Plates based on triangles, circles, etc. are simpler to identify while numbers introduce ambiguity or partial difficulties in recognizing the numeral, for example when only a few of the colour elements within the plate are confused (Lakowski, 1965b).

Essential structure of PIC tests

In the past the efficiency of PIC tests was evaluated by diagnostic criteria but now an objective theoretical approach is possible since colour measurement techniques and the theory of colour vision have advanced sufficiently for them to be confidently used. Although by applying spectro-photometric analysis we could specify these stimuli physically, only

colorimetric information will be used here since it is more akin to the relationships underlying stimuli and sensory discrimination. From such data *quantitative* predictions can be made about the difficulty of discriminating, say, figure and background in a given test. This is done by calculating Δ_C (or colour difference, see section on Colour specification), while a direct *qualitative* approach will allow us to determine whether or not colour confusion in PIC tests was adequately realized (as this is always deliberate), and this can provide an indication of how successful the given plates are in detecting outright dichromats.

The four basic plate designs common to all PIC tests are now described in greater detail with reference to colorimetric data. The names given to these plates are the ones first used by Hardy, Rand, and Rittler (1945).

Vanishing type of plates These are the simplest and most frequently used plates and are found in every PIC test, some being entirely composed of them. The colours of the figure and background are chosen so that they are aligned along a particular confusion line for a given dichromat. In dichotomous red-green tests the choice of coloured dots is made to fit both deutan and protan confusion loci so that neither is able to read the intended design. In modern tests such as the Tokyo Medical College (T.M.C.) colour vision tests they are used in the so-called 'screening' and 'quantitative' plates. Colorimetric data for these plates (see Fig. 1) and for subsequent figures are arranged so as to yield the maximum information. For example, (a) the chromaticities of coloured dots used in the tests can be read off to the second decimal point; (b) a number of isochromatic lines are selected in order to illustrate the colour confusions of a given dichromat to their best advantage; (c) the numerical values of Δ_C are marked in; and (d) by using the methods described in the section on 'Colour specification', the hues and saturations used either for each individual dot or for the total figure and background for each plate, can be found from λ_D and purity ratio.

The principle involved in designing vanishing plates is the same for all PIC tests so this example illustrates the structure of all such plates even though the chosen colours might be from a different area of the colour space. In this particular test the colours are chosen near the spectrum locus position so that the same coloured dots will serve for both protan and deutan defects since their isochromatic colours are almost the same at this point. The fact that colour differences are introduced permits a degree of quantification; Δ_C between the colours is quoted in Fig. 1 and varies from about 18 N.B.S. units for the screening plates to 30 for the right numeral of the quantitative plates and about 50

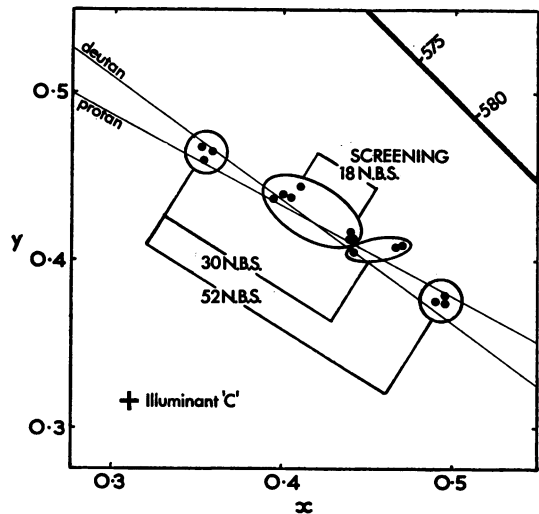


FIG. 1. Chromaticities for some T.M.C. plates, illustrating the 'vanishing type' of design.

N.B.S. units for the left numeral. This allows us to estimate the severity of defects, provided that none of the screening plates was read correctly. It is (a) mild or first degree if both numerals of the quantitative plates can be read, (b) medium or second degree if the left numeral can be read but not the right, and (c) severe or third degree if neither left nor right numerals can be read.

Qualitatively diagnostic type Basically these plates are an extension of the vanishing type of plate except that two clusters of coloured dots are used in printing *two* separate figures (usually one for the deuteranope and one for the protanope in the red-green type of plates, and one for the tritanope and one for the tetartanope in the yellow-blue type of plates). One cluster of colours is used for the background common to both figures. They were included in PIC tests so that the different types of dichromat could be distinguished, as essentially this is what PIC tests claim to be best able to do, that is, diagnose dichromats accurately. However, even here there are variations in design. For example, the Ishihara and Dvorine have much in common whereas the AO H-R-H and T.M.C. tests are differently designed. In some of the tests grey is used as a common background for both deutan and protan figures, but in the Dvorine it may be the figure that is grey and the background that is purple and red, while in the T.M.C. the deutan figure may be green and the protan red. It is very important that the two types of figures should be well placed along their confusion lines in relation to the background. The one described in detail here is a qualitatively diagnostic

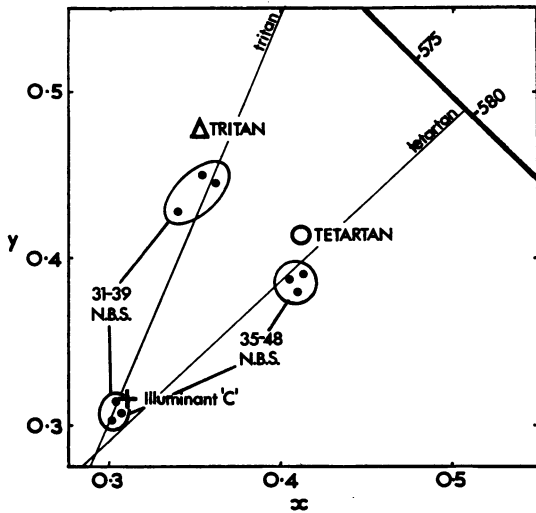


FIG. 2. Chromaticities for 'qualitatively diagnostic' plate no. 19 of the AO H-R-R test.

plate (no. 19) in the AO H-R-R test (see Fig. 2) designed to detect yellow-blue defects, which the normal observer sees as a yellow-green triangle and an orange circle. In relation to the grey background the colours of the triangle lie on tritan, and those of the circle on tetartan confusion lines, so that the tritan will see only the circle and the tetartan only the triangle. In this particular plate colour differences are quite large – about 35 N.B.S. units for the tritan figure and 42 for the tetartan.

There are other diagnostic plates in this test which utilize reds and purples to differentiate deutan from protan defects in much the same way as they are used in the Ishihara and Dvorine, but figures made from blues and blue-greens (for example in the AO H-R-R) and purples and greens (see the T.M.C.) could equally well be used. An examination of Fig. 6 in part 1 of this paper showing the confusion lines for different dichromats will indicate the various colour combinations theoretically possible but the further away we are from the straight part of spectrum locus in the yellow-green and orange area of the chromaticity diagram the easier it is to differentiate deutan from protan defects by means of surface colours. The same principle applies, though to a lesser extent, to yellow-blue defects except that here the line differentiating them moves from the purple to the blue-green area in the colour space. This aspect is not fully appreciated by some PIC test designers and therefore many of the diagnostic plates, though effectively dichotomous, do not distinguish between the different types of defect (*i.e.*, they are not good qualitatively diagnostic plates). On the whole,

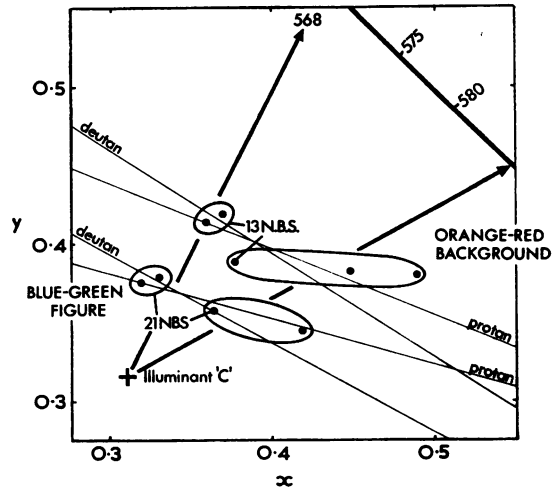


FIG. 3. Analysis of the 'transformation type' of plate – from Ishihara 5th edition.

ΔC is always largest in these plates (generally between 40 and 50 N.B.S. units), which ensures that only outright dichromats will make the expected misreadings.

Transformation type This is perhaps the most interesting and cleverly designed type of PIC plate, for both normal and defective observers can see a figure in these, though each identifies a different one. It is these plates that are the most intriguing. With colorimetric data and colour confusion theory we can understand how this is possible. At this point the reader should refer to Fig. 3 showing chromaticities for the Ishihara test plate no. 5 (5th edition) which is read as 74 by normal observers. Notice that there are four clusters of colours. The normal observer perceives a green figure on a reddish orange background, and the hue separation between them is about 15 nm. in terms of dominant wavelength notation (λ_D). To such an observer the reference point for colour is the position of illuminant 'C'. However, the reference points of the red-green dichromats are in the red or red-purple areas of the C.I.E. diagram and so their colour difficulties are different. If we follow their confusion lines, half of the normal background and half of the colours for the normal figure become the alternative 'figure', and the other half of the figure and background becomes the 'background' for the dichromat. This involves a rotation of almost 90° in colour space, altering colour 'affinities' appreciably, so that the normal observer sees a 74 (composed of green, light green, blue-green, and light blue-green dots) and the red-green dichromat sees the number 21 (made

up of green, brown, and orange dots against a background made up of blue-green, light brown, and pink dots). Δ_c between figure and background is about 40 to 50 N.B.S. units, except for two coloured dots in the background lying nearest to the greens in the figure where Δ_c is about 12 to 25 N.B.S. units. It is the proximity of these few elements in the plate that prevents people with small colour losses from seeing the intended pattern distinctly. An element of colour discrimination comes into play and so there is the possibility of reading other numerals, for example 1 instead of 4. Lakowski (1964, 1965b) found that the very young and older subjects find some such plates difficult to read 'correctly', and these can therefore be looked on not only as dichotomous plates but also to some extent as a measure of colour discrimination.

Hidden digit type These are plates specially designed for dichromats, who are expected to see a figure where normal observers should see none. This type of design is mainly found in the Ishihara plates, but one of Willmers' tritan plates (Stiles, 1952; Kalmus, 1955; Lakowski, 1966) is also based on this design. Figure 4 gives the chromaticities for the colours used in the Ishihara plates 10 and 11 (5th edition). In the previous three types we saw that colours for figure and background involved different hues which were usually separated by a large wavelength difference ($\Delta\lambda$). This is not true of hidden digit plates. Though three hues are used, there are variations in hue saturations and these prevent normal observers from forming a figure while allowing red-green dichromats to do so. The latter

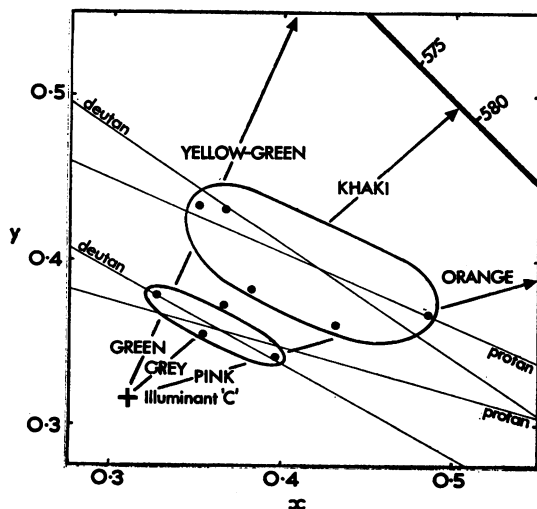


FIG. 4. Colorimetric data for 'hidden digit' plates. Ishihara 5th edition, plate nos. 10 and 11.

perceive two colour groupings, distinct enough from each other to follow two separate isochromatic lines, the more saturated orange, khaki, and yellow-green dots forming the background, and the less saturated pinks, greys, and greens making the figure. The readability of the hidden digit plates (in itself an error) depends on the degree of red-green defect as those whose defects are more extensive read these plates more frequently. The more recently designed PIC tests do not include such plates since some normal observers can also read them and this ability seems to be a function of age in normal observers. For example, up to 40% of subjects in the 20-30 age group read these plates easily, whereas they are hardly ever read by subjects over 50 or by young children (Lakowski, 1965b). This type of plate is found in all editions of the Ishihara but not in the Dvorine, T.M.C. or the AO H-R-R tests.

Contribution of objective analysis

Though the concepts of colour difference and colour confusion are useful for analysis, the results of such an analysis must be carefully interpreted and if the inferences are to be seen to be valid, ideally we should have confirmation from empirical studies. The possibility of classifying and quantifying defects from empirical data alone is to be examined in detail later, but at present the contribution of objective methods to our understanding of the PIC tests is now stated briefly.

Colorimetric analysis shows that broadly speaking the four tests - Dvorine, Ishihara, T.M.C. and AO H-R-R - conform to colour confusion theory since the elements in individual plates are aligned along the dichromat's confusion loci. The T.M.C. and AO H-R-R correspond more closely to these than do the Dvorine and Ishihara, but because the last two tests utilize a greater number of colours for the figure this may not be so critical.

We find that on the whole PIC tests use the central part of the C.I.E. space, utilizing less saturated colours than those used in tests based on the spectral principle. Confusions are arranged either between chromatic stimuli alone (Ishihara and partly also Dvorine) or between chromatic and neutral stimuli as found in the AO H-R-R and in the quantitatively diagnostic plates of the T.M.C.

In plates which were deliberately designed to classify dichromats, *i.e.*, the qualitatively diagnostic, we find that the correspondence to the appropriate confusion loci is again less close for the Dvorine and Ishihara than for the two newer tests, and this applies particularly to parts of the plates designed to detect protan defects. There is experimental evidence (Lakowski, 1964) to support this with reference to the 10th edition of the Ishihara and the 2nd edition of the Dvorine. In a group of outright deuterans and protans only one of 18 protanopes responded in

the accepted manner to the Ishihara, while three out of the 17 cases gave the intended reading for the Dvorine. However, of the 18 deutans 12 gave all the expected responses for the Ishihara qualitative plates and, similarly, of the 20 deutans tested on the Dvorine 10 gave the expected responses. Thus the percentages of 'expected' or 'correct' readings of the possible totals is 22% for protans and 88% for deutans in the Ishihara, and 20% for protans and 60% for deutans in the Dvorine. This deficiency, of course, does not detract from the fact that these plates are still excellent dichotomous tests, and as there are colorimetric differences in the various editions of such PIC tests, these findings may not apply to other editions.

We have also seen that colour difference between the stimuli of figure and background is always largest in the qualitatively diagnostic plates and less for the 'quantitative' or 'screening' plates. With reference to Δ_c measurements on the AO H-R-R the validity and usefulness of the concept of colour difference for assessing the difficulty of a colour task can be illustrated. It was found that for the screening plates Δ_c is 15 to 25 N.B.S. units, while for plates designed to detect medium R-G defect it is ≥ 36 N.B.S. units and for plates for detecting extreme R-G defect it is ≥ 42 N.B.S. units. Here we have an example of a test designed to detect *degrees of defect* where the increased Δ_c found objectively corresponds to what the designer intended. Similarly, in the T.M.C. screening plates, Δ_c is 18 N.B.S. units while in plates for detecting mild, medium, and severe defects it is 18, 30, and 50 N.B.S. units respectively.

In these two tests then, the order of presentation of the plates is according to the planned order of difficulty. This does not apply to the Ishihara and Dvorine tests as these were designed mainly to be dichotomous and therefore all plates should be equally difficult. However, it is interesting to note that the colorimetric data reveal that though in most of the plates, for example the Ishihara, Δ_c is 25 to 45 N.B.S. units, in some, such as the transformation type (series I and II) which contain elements critical to the perception of the arabic figure, it is only between 10 and 15 N.B.S. units, and in such cases it is better to consider the plates as useful for screening or detecting slight defects. The same applies to the Dvorine where Δ_c is only about 10 to 15 N.B.S. units for the plates in series V (plates 10 and 11), and about 18 to 20 N.B.S. units for plates in series IV and there is empirical evidence to show that plate 10 is 'difficult' to read. It has the highest percentage of misreadings per plate in this test, that is, almost 40% in a population of 500 subjects (Lakowski, 1964). Again these plates could be considered screening or quantitative plates for detecting mild defects or losses in discrimination.

So, by having the appropriate knowledge, we could rearrange the order of presenting the individual plates in such dichotomous tests where those with small colour differences (say less than 20 N.B.S. units) could be used for detecting minor defects or impairment of discrimination in normal subjects, while those with larger differences (35 to 40 N.B.S. units) would be used for detecting outright dichromats.

The problem of classification

We have seen that the aim of all PIC tests is to detect colour confusion. If there were no individual variations in defects and if the tests were perfectly designed, *one* plate alone would suffice for that purpose. But the ideal plate has yet to be designed and we must also accept that there are degrees of defect. The poor discrimination of some extreme anomalous trichromats which can amount to partial colour confusion may cover large areas of the colour space. Such subjects may perform like outright dichromats, and so we find that the problems of discrimination and confusion are intermingled. Many tests are too crude to distinguish between the large difference in defects (see section on Anomaloscopes, Fig. 15). In the older PIC tests the aim was to divide the population only into those with and those without congenitally determined red-green colour defects (dichotomy), whereas the more recent tests have introduced discrete steps, so that some 'quantification' should be possible. The problem of diagnosis is further complicated because there are subjects who, though not congenitally red-green deficient, still give faulty readings mostly for the vanishing type of plates and occasionally for the transformation plates.

How far then can PIC tests be used to *measure* colour vision and can the degree of defect be quantified from the number of misreadings? The T.M.C. and AO H-R-R have ostensibly done this already, but what about the other tests? Data quoted here are confined to the Ishihara and Dvorine, which are examined to find whether quantitative diagnosis is possible, and as it is difficult to make absolutely final classifications we shall discuss diagnoses only in terms of probabilities.

First let us examine a hypothetical scale of probabilities that could apply to the Ishihara readings. If a subject misreads 24 of the total 24 plates given, the probability that he has a *congenital* red-green defect is about 0.9 (that is, very high) though this cannot be absolutely certain. His faulty reading might, for example, be the result of central scotoma, such as is found in the early stages of *acquired* dyschromatopsia due to alcohol amblyopia. That he is an outright dichromat is less certain (with p about 0.80) for extreme anomalous subjects can score very large numbers of misreadings. If he

gives the expected deutan readings the probability that he is a deutan is 0.85, but the prediction for the classical protan type of defect is less certain as many such subjects do not give the expected type of misreadings, for example in the 10th edition of the Ishihara (Lakowski, 1965b, 1966). On the other hand, with such numbers of misreadings, the probability that the subject is a simple deuteranomalous is small (thus $P = 0.2$) and that he has normal colour vision is even smaller ($P = 0.05$), though there is always a possibility that a neurotic person with no congenital colour defect could obtain such a high score (Pickford, 1949). Of course, the probability of correct diagnosis will be increased enormously if information available from other colour vision tests, especially from tests based on a different principle, is used.

Certain research data (Pickford, 1950; Belcher, Greenshields, and Wright, 1958) appear to show that the frequency distribution of misreadings of, say, the Ishihara plates in a randomly selected population has almost no continuity between normal and colour defective subjects. Such frequency distributions are J-shaped with a large number of subjects making a few mistakes at one end of the error scale and a relatively large number of subjects making many mistakes at the other end of the error scale. But it should be remembered that these data were collected on young subjects between 20 and 30, many of them students, who socially and intellectually are scarcely a representative sample of the general population. For such a population the division between the number of misreadings made by defective as opposed to normal subjects is easily established, but this is not necessarily the case for others, where the dividing line might depend very much on the sampling method, on age, or even on what constitutes a 'misreading'. Let us consider how a given defect is characterized in terms of numbers of incorrect readings. Classification of defect is accord-

ing to that obtained from the P-N anomaloscope where the term 'outright dichromat' denotes those deuteranopes (D) and protanopes (P) who easily accept all the possible ratios in the Rayleigh equation. Simple anomalous trichromats are those deuteranomalous (DA) and protanomalous (PA) subjects who deviate significantly in their matching ratios from those most frequently accepted by normal observers, but otherwise have good discrimination, while the term 'incomplete' implies all subjects not included in the above classifications EDA and EPA, that is anomalous trichromats with enlarged matching ranges (accepting a number of ratios) and incomplete dichromats (see section on Anomaloscopes).

Number of subjects (all males) in each sub-class are quoted along with the highest and lowest scores obtained (see Table 1). The data are for the Ishihara and Dvorine plates only as data at the Visual Laboratory are more extensive for these tests than for any other PIC tests.

In the first instance diagnosis from Dvorine and Ishihara results could be according to the mean scores which are lower among those with less severe anomalies. Variance is small in the scores of outright dichromats but much greater among the least defective (PA and DA). Results indicate that high error scores can be found with all degrees of defects, while low scores, similar to the scores of normal observers, are found only among the simple anomalous. However, the very low scores of the minor colour defectives do not occur often, and therefore the data might exaggerate the extent of any overlap. By giving data for only 70% of the population sample (where the top 15% and the bottom 15% are excluded) we arrive at a more realistic picture of the efficacy of the tests. In Figs 5 and 6 the two percentile limits are quoted along with the mean data shown as curves to make the picture more comprehensible (no mathematical properties

TABLE 1
DISTRIBUTION OF INCORRECT READING OF RED-GREEN DEFECTIVES FOR TWO PIC TESTS

PIC test	Outright dichromats		Incomplete		Simple anomalous trichromats	
	Protan	Deutan	Protan	Deutan	PA	DA
Ishihara misreadings	<i>N</i> = 21	<i>N</i> = 23	<i>N</i> = 13	<i>N</i> = 46	<i>N</i> = 6	<i>N</i> = 19
Highest	24	24	24	24	22	22
Mean	22	22	16	19	7	12
Lowest	20	18	2	6	1	1
Dvorine misreadings	<i>N</i> = 20	<i>N</i> = 22	<i>N</i> = 14	<i>N</i> = 47	<i>N</i> = 4	<i>N</i> = 19
Highest	14	14	14	14	14	14
Mean	14	13.8	10	12	4	8
Lowest	14	13	3	3	0	0

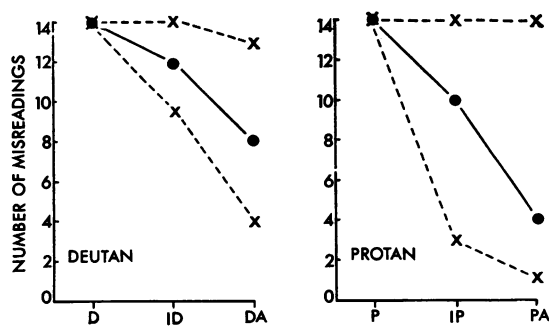


FIG. 5. Dvorine misreadings for the mean, 15th, and 85th percentiles for deutan and protan type of defects; dichromats (P, D), incomplete (IP, ID), anomalous (PA, DA).

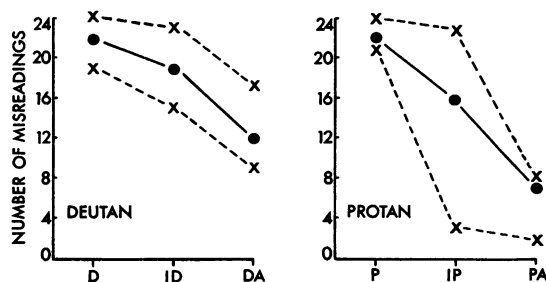


FIG. 6. Ishihara misreadings for the mean, 15th, and 85th percentiles for deutan and protan types of defects; dichromats (P, D), incomplete (IP, ID), anomalous (PA, DA).

should be attributed to these). Bearing in mind that the data on anomalous deutan defects are drawn from a larger sample than was available for protans, the inferences that can be made from these two diagrams are numerous, for example both Dvorine and Ishihara discriminate degrees of deutan defect better than comparable protan ones.

In order to find the highest possible number of misreadings for non-defective observers let us look at *low* scores alone. Would such an approach allow us to exclude colour defectives, and how certain can we be that it does so? The usual criterion (Belcher *et al.*, 1958) used for normal observers is that, on first presentation, there should be not more than three misreadings for the Ishihara and not more than one for the Dvorine. It must be emphasized, however, that these are very biased figures which could be misleading if classification of the results of the very young were made according to these criteria, or if it were used on subjects older than those involved in Belcher's study. If we look at Fig. 7 giving average misreadings of a modified age

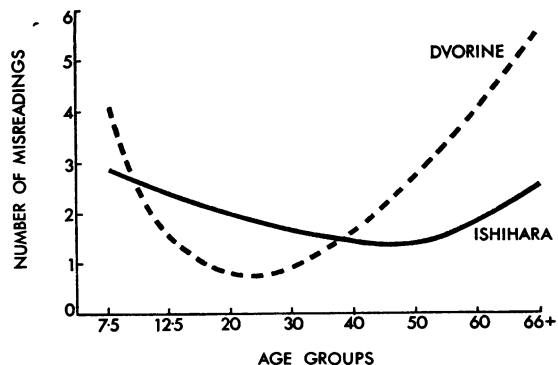


FIG. 7. Mean number of misreadings for an age population on two PIC tests.

population for the two PIC tests, we see that the cut-off point used by Belcher *et al.* could only apply to certain age groups. Thus the criterion used to decide the cut-off point depends not only on the type of population involved but also on the sampling technique employed. Let us consider this last point in greater detail. We distinguish between two kinds of population (*a*) *inclusive*, where everyone is tested including all those with red-green defects, and (*b*) *modified*, where colour defective persons are excluded from the sample. Again we may have a 'closed population' where everyone in a given class or locality is included and tested, or an 'open population' where some choice has been made and thus only a 'sample' of a given class is tested. In such populations it is important to assess the extent of the 'randomness' achieved. For this purpose Belcher's results will be compared with those obtained from 745 male police candidates (Peters, 1954), with those of 364 first-year psychology students at Edinburgh, and with an age population of 500 subjects – 244 male and 256 females – ranging from 5 to 80 years (Lakowski, 1964). Analysis is confined to data for the Dvorine, as discrepancies are largest here, though there are similar smaller ones for the Ishihara. Figure 8 shows results for the four *modified* populations (that is, red-green defectives excluded). The frequency curves for three of the populations show a less sharp decrease in numbers of subjects with greater error scores than were found in Belcher's population, where a great many subjects made only one or two misreadings. The frequency curve for the age population is quite different from the other three in having the lowest percentage of subjects with no misreadings, and in showing a 14.5% incidence of subjects with seven or more misreadings. Remember that both the Edinburgh student study and Peter's study were on *closed* populations. Both these studies show high percentages of those with two and more misreadings (32 and 16 respectively). In the age

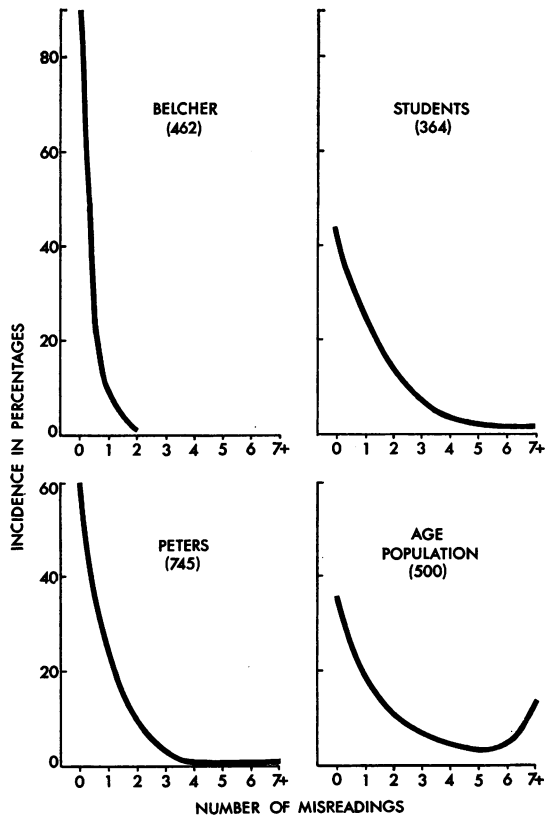


FIG. 8. Frequency distribution of misreadings on the Dvorine for four modified populations (red-green defectives excluded).

population, half of the subjects would be above Belcher's criterion and therefore would be suspected of having defective colour vision.

It is useful to examine the distributions of misreadings in the age population in greater detail (see Fig. 9 showing four sub-groups each involving about 120 subjects). Only in one age group (16-36) does the frequency distribution resemble a reversed J type of curve similar to Belcher's results. Thus if Belcher's cut-off criteria are accepted, a glance at the frequencies shown in Figs 7 and 9 will indicate that a large number of subjects would be suspected of being colour defective, especially if the results were only on young subjects (*e.g.*, in vocational guidance cases). Or if older subjects alone were being considered (*e.g.*, for up-grading or retirement) as many as half of them would have to be retested by other means. It appears therefore that the criterion used is of doubtful validity and so a more flexible approach becomes necessary. The dividing line must be made at a higher level of misreadings, but, as we have already seen, this would overlap the error scores of congenital red-green defectives. This further confirms that the cut-off point between normals and defectives is not easily made and that when dealing with low level scores we are again dealing only in probabilities. Thus if a subject makes seven misreadings on the Ishihara and, say, four on the Dvorine, and is between 20 and 35 years of age, the probability is $P = 0.7$ that he might be a PA or DA, but with a decrease or increase in errors the probability of his having in the first case normal and in the second case defective colour vision becomes greater.

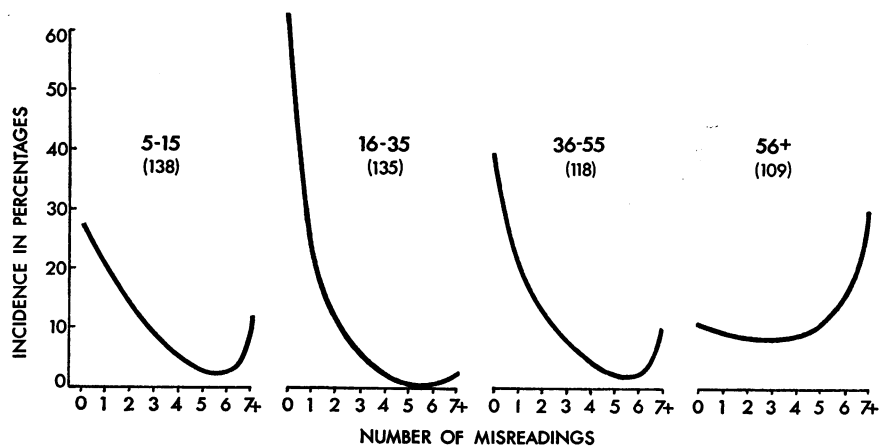


FIG. 9. Frequency distribution of misreadings on the Dvorine for four age groups (modified).

In brief, it is evident from what has been said that it is not possible to ascertain the degree of defect with any certainty from the number of misreadings made on the Dvorine or Ishihara, though it is claimed that this is possible with the T.M.C. and AO H-R-R tests (Umazume and Harutake, 1962; Hardy *et al.*, 1945). This is because those with minor congenital red-green defects (anomalous trichromats) can make either a very large or very small number of misreadings. A final diagnosis of the extent of the defect based on misreadings of PIC plates alone is always difficult, never certain, and at best only probable. In the case of outright dichromats the difficulties are fewer though some are experienced with the qualitative plates.

Colour discrimination – the 100-Hue test

The Farnsworth Munsell 100-Hue test, which is based on recognition of surfaces by reflection, is a test of colour discrimination and colour confusion designed by the late Commander Dean Farnsworth (1943, 1957) in the early forties. It includes many features not found in PIC tests, for in addition to detecting colour confusion it can indicate minute differences in colour discrimination. The test consists of a scoring sheet and four wooden boxes that together hold 85 movable caps, in which Munsell colours are mounted. As the colour task involves arranging these caps in a continuous 'colour series' it was soon found necessary to subdivide this large number of stimuli into smaller task units. The caps were divided into four groups of about 21 and placed in wooden boxes, and the arbitrarily chosen groups became the four series of the test, one red to yellow, a second from yellow to blue-green, a third blue-green to blue, and the fourth from blue to purple-red. The different coloured caps each numbered on the back can be moved about freely during the performance. Each box also contains two extra immobile caps, the first and last of the neighbouring series (the positions of these in the C.I.E. space are shown in Fig. 10).

By selecting Munsell colours equidistant from illuminant 'C', the resulting cap colours are of equal saturation, and brightness control has been established by maintaining a more or less constant level of luminance for each of the 85 caps (in Munsell notation Chroma 6 and Value 6). As a result hue discrimination for surface colour is the only variable being tested.

Colorimetric analysis (Lakowski, 1966) shows that the colour difference (ΔC) between the caps is very small indeed (much smaller even than the smallest differences found in PIC tests), ranging from 0.6 to 5.7 N.B.S. units, with a mean ΔC between the caps of about 2.2 N.B.S. units. Colour differences for each box are not uniform, the box with caps 85 to 21

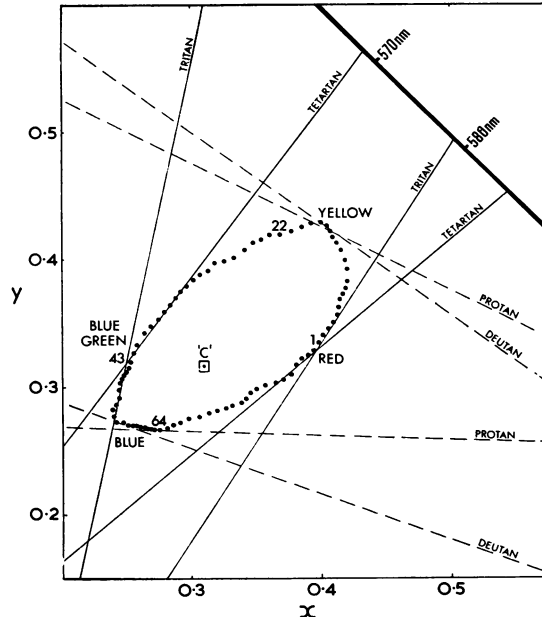


FIG. 10. Chromaticities for the 85 100-Hue caps. The tangential isochromatic lines show areas where dichromats have poor discrimination.

being the least difficult and the box with caps 43 to 63 the most difficult.

One of the great merits of the 100-Hue test is that those elements suitable for detecting small variations in colour discrimination are also suitable for detecting colour confusion. As the coloured caps were chosen to cover the entire colour circle it happens that in some areas certain consecutive caps follow the confusion lines of all known dichromats. Figure 10 shows the chromaticities of the 85 coloured caps, within the central part of the C.I.E. diagram forming an oval outline (it would be in the form of a circle if uniform chromaticity space was used). These chromaticities are from measurements made at the visual laboratory. Four isochromatic pairs of lines are also drawn, tangentially to this ellipse, showing approximately which of the caps are confused by deuterans and protans, and which by tritans and tetartans. In addition to detecting classical types of dichromats, Verriest (1964) has recently shown that the 100-Hue can detect the so-called scotopic type of confusion characteristic of many subjects with acquired colour defects.

The validity and reliability of the test for detecting the three well-known congenital types of dichromat are high and correlate well with similar findings on the Pickford anomaloscope and with data obtained from the König-Helmholtz colorimeter and also with results obtained from dichotomous tests such as the Ishihara (Lakowski, 1968d).

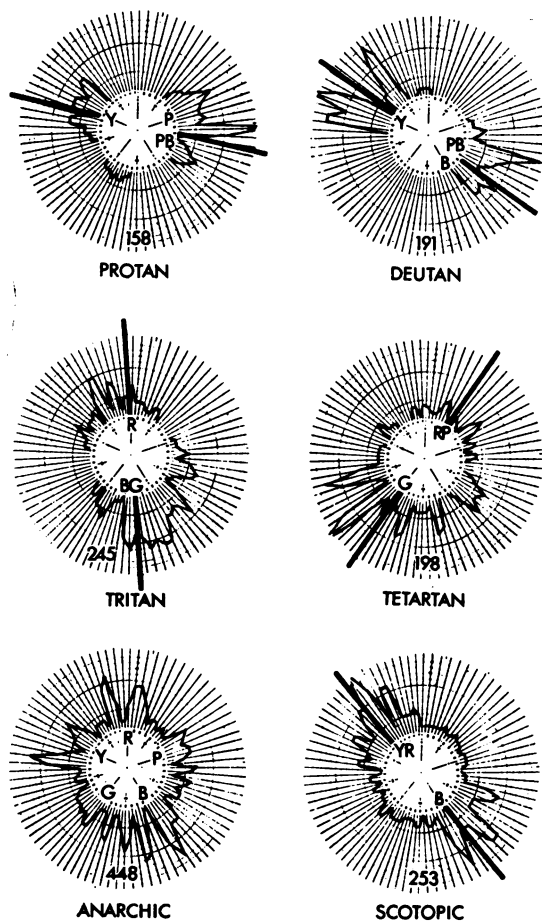


FIG. 11. Representative configurations for dichromats and the anarchic and scotopic patterns, showing characteristic axes, profiles, and total error scores for each type of defect.

Reliable and valid results can only be achieved when the testing procedure is standardized and is similar in each testing situation. It is recommended that the illumination should be of type 'C' or 'D' in the C.I.E. system and the intensity should be about 300 lux or 25-30 lumens/ft².

For normal subjects and for industrial populations Farnsworth suggested a time limit of two minutes for the presentation of each box, but this has to be altered in clinical situations or with subjects whose intelligence is impaired where it is not advisable to set a time limit. The test cannot be used with children in its present form but modifications were suggested in the form of a 40-Hue test, and a 28-Hue test by Ohta (1966) and Roth (1966) respectively. The individual caps in each box must be shown in random order and it is good practice to stick to the same chosen random presentation. The subject must

place, within each box, the individual caps in a 'colour series', that is, in such a way as to form a consecutive colour order made up of the individual caps.

Profiles and norms for the 100-Hue

To aid scoring, spaces are provided on the data sheet for the examinees' order of arrangement. If this order is the same as the printed one there is no error score, but, if not, the extent of transposition is calculated. The error score for a cap is the sum of the differences between the number of that cap and the numbers of the chosen adjacent caps. A score of 2 does not constitute an error, as this is a minimal difference. In this manner the so-called *partial error score* for each cap is arrived at, and additions of these make up the *total error score*. When the total error score is large it is helpful to draw out a *profile* and there are various methods of treating the results, depending on the information required. The pattern for dichromats is identified by *polarity*, that is clustering of maximum displacements (errors) in two regions which are almost opposite forming the so-called axis of defect. Figure 11 shows six representative patterns for such dichromats with clearly marked axes for each defect. Other configurations can be found in François and Verriest (1961), Cox (1961a), and Lakowski (1968a).

The letters within the centre part of each pattern show the regions of the test where discrimination is impaired. For example, for protans this is the yellow, purple, and purple-blue regions, while for tritans it is the reds and blue-green parts. The particular colour names involved in the regions of poor discrimination in this test must not be confused with the pairs of colours used to show classical dichromat colour confusion; that there is no contradiction here can easily be verified by consulting Fig. 10 again. We can now be more specific and state which of the cap numbers are likely to be not discriminated by dichromats. These are caps 14 to 24 and 57 to 72 for protans, caps 12 to 22 and 52 to 64 for deutans, caps 80 to 9 and 42 to 54 for tritans, and caps 74 to 1 and 30 to 44 for tetartans.

It is also possible to calculate the total error score and this is a more useful method of assessment in defects that are less than outright dichromatism.

The test is a relative measure of colour discrimination and therefore certain criteria or reference standards are necessary against which to evaluate the given individual score or configuration. Some of the norms are found in Table 2, where median values and various percentile scores are given for Farnsworth's industrial population (1957). In addition norms collected at the Visual Laboratory in Edinburgh for other populations are quoted. These include two non-clinical populations. All the relevant information about these groups is given in

TABLE 2
NORMS FOR THE 100-HUE

Statistical parameters	Farnsworth's norms (1)		Random sample (2)	Printers' apprentices (3)	Diabetic patients (4)	Arthritic on chloroquine (5)
	Unselected	Industrial				
No. of subjects (N)	—	300	57	106	507	59
Mean	—	32	55	70	123	132
Standard deviation (σ)	—	29	40	36	71	75
Percentile scores						
Median— 20th	20	8-12	15	40	65	65
50th	46	24	45	60	112	115
70th	76	36	65	85	148	165
90th	120	68	125	125	213	255
95th	—	—	155	140	245	285

(1) from the Manual, 'unselected' group 15-45 years of age; 'industrial' refers to in-plant applicants in a paint factory. From Visual Laboratory; (2) Edinburgh age sample, showing agreement with Farnsworth's data; (3) printers' apprentices, young boys 13 to 16 years of age; (4) diabetic population - monocular testing, age span 10-70 years; (5) rheumatoid arthritic patients on chloroquine, 59 eyes, 20 to 50 years of age.*

*Lakowski, R., Haining, W., and Partridge, R. (1968). Colour Vision Losses in Arthritic Patients in Early Stages of Chloroquine Therapy. Preliminary Report. Preprint from the Visual Laboratory, University of Edinburgh, No. 31/2, March 1968.

the table. It was considered necessary to quote scores for these various statistical parameters, since, as with all data on sensory discrimination, the distribution is rather skewed or of 'J' type, and as a result values for the means and standard deviations alone might be misleading, but these are quoted here so that comparisons can be made with the next set of norms obtained for their age population by Verriest, Vandevyvere, and Vanderdonck (1962). Here mean scores and sigmas are given for six groups from 10 to 60 years of age.

TABLE 3
VERRIEST'S AGE POPULATION IN DECADES
(AFTER KINNEAR, 1965)

Mean age	N	Mean score	σ	Scores		95th Percentile
				Lowest	Highest	
10s	56	51.5	28	8	124	100
20s	145	40	31	4	162	92
30s	70	56	34	8	176	120
40s	62	76	37	16	184	144
50s	69	83	38	12	176	164
60s	29	88	35	28	174	152

Although configurations for outright dichromats are well defined, this does not hold good for those with lesser defects, or for the variations found with normal trichromatic discrimination. For such subjects the cut off in terms of raw scores is always an arbitrary one. We may want to divide subjects into two sub-populations, those with scores above the

median and those with scores below it, or we might be interested in selecting only those with the best performance in which case we would choose people with scores at about the 20th percentile, or, on the other hand, we might only wish to eliminate those with poor discrimination, say those with scores around or above the 75th percentile. Data for the 95th percentile are included whenever possible as this is a useful statistical parameter which indicates what the odds are against a large score appearing by chance in a given population, and statistically such chance ratio (1:20) is the minimum acceptable significance level for a given individual score ($p = 0.05$ level). From the table we see that the 95th percentile score for the 100-Hue in the non-clinical population is equivalent to scores at the 70th percentile for diabetics. We may infer therefore that in a diabetic population the chance of obtaining a similar high score on the test is 3 in 10, which is significantly different from that of the non-clinical population.

Thus the 100-Hue test is a useful test of colour vision. It detects all forms of dichromatism even of the yellow-blue type, and is very useful for detecting minor deviations in colour discrimination in normal trichromatic observers and an excellent means of identifying acquired colour discrimination losses. It is relatively quickly administered, taking about 15 to 25 minutes to complete. Perhaps its major drawback is that it does not identify anomalous defects of the DA and PA type, but this cannot be considered to be a valid criticism since it was never designed to do so.

Metamerism – anomaloscopes

In their comparative evaluation of anomaloscopes Willis and Farnsworth (1952) define these instruments as devices 'for mixing red and green lights in varying proportions to produce a match for yellow light . . .'. This is true of the instrument used by Lord Rayleigh (1881) who first discovered the wide variations in the ratios of subjects with apparently normal colour vision. Since then, and especially since the construction of Nagel's anomaloscope in 1898, they have been used to find anomalous and dichromatic red-green defects (Nagel, 1907).

Essentially anomaloscopes are instruments, where either spectral or filtered lights employing a film mode of appearance are used, in which the viewing aperture is either fixed or can be varied slightly. The viewing field is divided into two parts, one standard and one variable, the latter involving a mixture of two primaries. Matches producing an equal sensation for the two parts are in a metameric relationship, and so in the original 'Rayleigh Equation' a variable mixture of lithium red and thallium green were matched against a sodium yellow. In some of the Nagel anomaloscopes (model II) the so-called 'Trendelenburg Equation', in which a standard blue is matched by additive mixtures of violet and blue-green, is included. Finally, the Pickford type of anomaloscope utilized a third equation in which a white (near illuminant A) can be matched by a mixture of yellow and blue (see Fig. 12).

The information that can be obtained from an anomaloscope will depend on the method of testing and on how the data were quantified. There are a variety of psychophysical methods of assessing colour vision by this means, and in the early days a subject was permitted to operate the variables himself, yielding one matching setting. This method was sometimes extended by allowing the testee to make several matches and a mean setting was calculated from this. Others recommended that all manipulations should be done by the experimenter and that the method of constant stimuli should be utilized. Pickford (1951), however, strongly recommends the use of the methods of limits, where the examiner makes a number of serial adjustments, continuing until he has established all possible mixture ratios and explored the whole matching range or interval of equality for the observer.

Some subjects might have a small *matching range* (MR), accepting only one ratio setting, whereas others may accept a number of ratios and thus have a larger MR, and when this happens a *mid-matching point* (MMP) is usually calculated. From the combined data of matching range and mid-matching setting inferences can be made about the sensory discrimination and physiology of a subject's visual system.

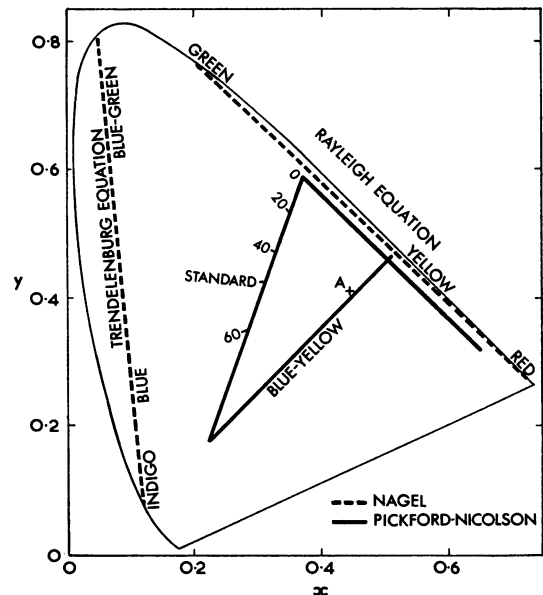


FIG. 12. Loci of anomaloscope colour equations giving their original names, and chromaticities for the green-blue equations 0, 20, 30, and 60 arbitrary unit settings and standard (for the P-N anomaloscope).

There are a number of methods of recording such data, the simplest being to quote the mixtures accepted, a kind of arbitrary instrumental index indicating the numbers of millimetres or degree settings on a scale. Another way is to mark such settings on a diagram against an arbitrarily chosen scale representing the entire colour equation – with its red and green end points.

Deviations in mixture ratios

Though a large number of subjects have very small MRs and accept perhaps only one mixture ratio so that both their MMPs and MRs have the same numerical value, there are others, who, although their discrimination is good, require differing amounts of the primaries to match a given standard, some needing more red and others more green in the Rayleigh equation. For such people it is both useful and meaningful to express this deviation quantitatively. In order to do this Trendelenburg (1929) introduced the so-called 'anomaly quotients' for the Nagel anomaloscope, a mathematical device which attempted to express the magnitude of such deviations by the following formula:

$$Q = \frac{N'}{73 - N'} \div \frac{N}{73 - N}$$

where N is the position of the mixture setting for an average normal observer, and N' the position

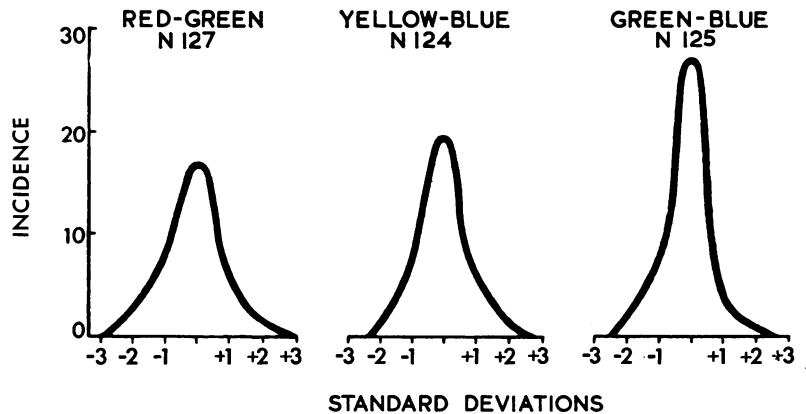


FIG. 13. Frequency distribution of mixture ratios for three P-N anomaloscope equations (red-green, yellow-blue, and green-blue) and number of subjects involved (age span 16-60 years).

adopted by a deviant subject, and the number 73 represents the end of the scale for that particular Nagel instrument in the red-green equation. The more the value of Q deviates from 1 in either direction the greater is the deviation from the most frequent quotient. For example, $Q = 1.5$ denotes a degree of deuteranomaly, while $Q = 0.5$ indicates a degree of protanomaly, and any intermediate values can be obtained. This measure is most frequently used on the continent, and is a very simple way of expressing quantitatively certain aspects of, say, red-green anomalous vision. Verriest (1968) shows that this formula can be adapted for use with any type of anomaloscope.

The criticism of this is that it is still only a relative and highly arbitrary method of indicating deviations, meaningful only if we first know empirically which of the quotients denote anomaly. It is also a less sensitive measure of minor deviations. To avoid such criticisms Pickford uses statistical methods to measure the degree of deviation by relating each score to the central tendencies within a population and by expressing deviations in terms of so many standard deviations from the mean values. This is a valid method since in a random population matching ratios are normally distributed around a modal value (see Fig. 13, where these are given for the red-green, yellow-blue, and green-blue equations for about 120 subjects of both sexes, whose ages ranged from 16 to 60 years), and such treatment immediately refers the individual MMPs to a population distribution and so provides a more meaningful measure. Thus, one can talk about a given mid-matching value on the anomaloscope as falling within one, two, etc. standard deviations from the mean, indicating what the chances are of such a deviation occurring in a normal population.

For example, Pickford (1951) considers subjects with MMPs outside $\pm 2 \times \sigma$ as belonging to a type of colour aberration which he calls *deviant*. Thus we can talk about *red deviants* or *green deviants*. The limits of $\pm 3 \times \sigma$ mark the boundaries of normal distribution, and outside these limits lie MMPs of genetically determined anomalous trichromats.

Interval of equality or matching range

Pickford (1951) also attempted to quantify the discrimination date provided by the anomaloscope *matching ranges*. As the distribution of these is highly skewed or J shaped in a random population (see Fig. 14), parametric statistics are not applicable and he used the modal values instead, suggesting that subjects with two or more times of the modal MR should be called *colour weak* (with or without marked deviations of MMPs). Lakowski (1968d) suggests the use of percentile scores as a useful measure of colour discrimination. Here the mean and percentile values of MRs, especially with the 95th percentile placing, give us a direct comparison with population statistics and their incidence in a given population. These become very useful and are necessary for evaluating large ranges, for example in studies of acquired colour discrimination losses.

By applying spectrophotometric and colorimetric analysis, Lakowski (1965a, 1968d) has recently shown that anomaloscope results can be expressed in objective terms, giving not only the end points but also any intermediate positions for the various arbitrary anomaloscope settings in dominant wavelength (λ_D) or by quoting chromaticities within the C.I.E. diagram; the positions, for example, of the 0, 20, 40, and 60 arbitrary unit settings and the standard of the green-blue equation for the P-N anomaloscope are marked in Figure 12.

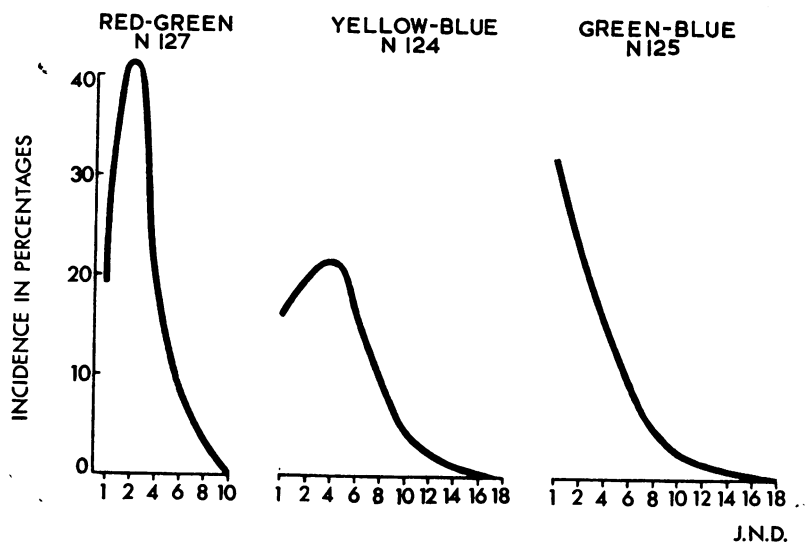


FIG. 14. Frequency distribution of matching ranges (discrimination) for three equations plotted against j.n.d. scale.

RED-GREEN COLOUR EQUATION	CLASSIFICATION	
	NORMAL	
	RED-GREEN DEVIANTS	VARIATIONS IN NORMAL TRICHROMATS
	COLOUR WEAK	
	SIMPLE PA AND DA	ANOMALOUS TRICHROMATS
	ENLARGED PA AND DA	
	EPA AND EDA	
	INCOMPLETE P AND D.	DICHROMATS
	PROTAN AND DEUTAN	
-3x ±S.D. +3x	STATISTICAL PARAMETERS	
0.5 1.0 1.5	ANOMALY QUOTIENTS	

FIG. 15. Graphical representation of red-green colour vision classifications yielded by anomaloscope data (see text for fuller description).

In addition, deviations in an observer's matching points or the extent of a matching range can now be expressed in 'just noticeable differences' (j.n.d.) as calculated from Judd's uniform chromaticity diagram (1935). This method of analysing anomaloscope data permits arbitrary deviations to be expressed along a uniform scale of colour sensation.

Figure 15 represents pictorially the type of information that can be obtained from anomaloscope data, showing the limits of the normal variations in MMP, deviant, and colour weak vision, and also some examples of congenital defective performances. In it the two central vertical lines represent $\pm \sigma$ limits, that is the variation of MMPs for 68.3% of the subjects in a normal population. If this is extended to include the outer lines it represents 99.74% of the variance or $\pm 3 \times \sigma$. The horizontal black lines cutting across or lying within these boundaries represent the MR of the specified types of subjects. In the first two groups classified, the first of the normal subjects and the four red and green deviants have good discrimination, accepting only one mixture ratio setting while the four colour weak subjects accept a greater number of ratios: they simply have large MRs (at least twice the median value of the MR of the normal trichomats). Deviants differ from the normal subjects in that their MMPs are at or outside the limits of $2 \times \sigma$ (not marked in the Figure). The same applies to the MMPs of the simple DA and PA subjects (protanomalous and deuteranomalous) whose discrimination on the anomaloscope is as good as that of normal observers except that their mixture ratios are well outside the limits of $\pm 3 \times \sigma$. On the other hand, those PA and DA subjects with enlarged MRs have poorer discrimination since they accept a greater number of mixture ratios. Note that what distinguishes the anomalous from all other subjects is that their MMPs are always well outside the limits of the distributions for normal observers even when one of the end points of their MR cuts across the position of the mean MMP of such observers. Dichromats, however, have MRs that extend through all the available mixture ratios for the given equation. Thus they accept all mixture ratios as equivalent to the standard whether they are the ratios typical of normals, anomalous or any other type of subject.

Whichever way the results are recorded or analysed the anomaloscope is an instrument capable of measuring continuous variations in colour vision, and in this sense it is far superior to any of the tests described here. Nevertheless it is a simpler device than say Wright's or the König-Helmholtz colorimeters, in which more complex mixture ratios can be chosen, and which have a greater degree of spectral purity. Because they are optically very complex such instruments require a greater degree

of sophistication on the part of both the experimenter and the observers, and of course they are more expensive and thus inappropriate for routine colour vision testing, although they are excellent research instruments.

There are a number of anomaloscopes at present in use, mostly manufactured in the United States. As few are obtainable in this country only two models which are easily available are described in detail – the Nagel anomaloscope manufactured by Smith and Haensch of Berlin and the Pickford-Nicolson anomaloscope produced by Rayner of Rayner and Keiller Ltd. of London.

The Nagel anomaloscope

Two models of the anomaloscope are produced; in model I only the Rayleigh equation can be utilized, while in model II the Trendelenburg equation is added. The field is circular with a horizontal division, and an adjustable diaphragm giving the following angles $1^\circ 15'$, $2^\circ 10'$, and $3^\circ 15'$ (in practice the last aperture is the most frequently used); the retinal illumination is at the order of 250 trolands (see section on Administration). The optics of the instrument are illustrated in Figure 16. In the upper field a yellow radiation of 589.3 nm. is shown and luminance in this part of the field can be varied by

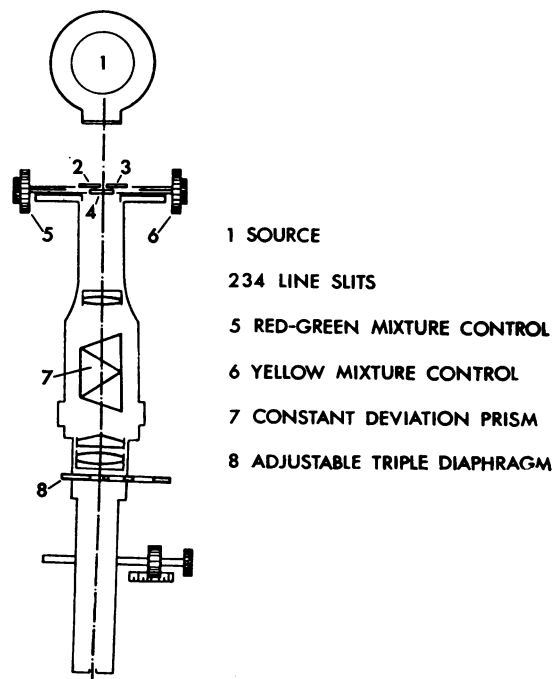


FIG. 16. The Nagel anomaloscope.

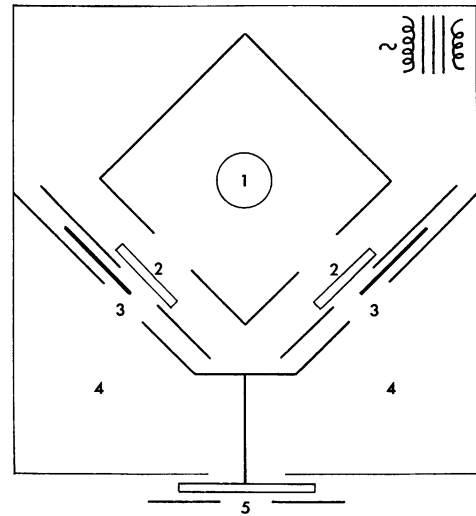
a control screw. The lower slits give the additive mixtures 670 nm. and 536 nm. In model II indigo and blue-green are matched to blue, that is a 490 nm. standard is matched by 470 nm. and 517 nm. In this model tritanopes and anomalous tritanopes can be detected where the dichromats equalize the indigo and blue-green with the intermediate wavelength while tritanomalous subjects use too much indigo for their match. Again it is possible to calculate anomaly quotients from the positions of the left-hand knobs exactly as for the Rayleigh equation. Because the subject's modal values vary greatly and because the saturation cannot be altered the equation is difficult to use and recent models of the Nagel anomaloscope have been produced without the Trendelenburg equation.

Of course it is important to have the instrument calibrated from time to time especially when it is first used after being out of use for some time, or after being moved about. This can easily be done with the help of sources with line spectra.

The Pickford-Nicolson anomaloscope (P-N)

From the optical and mechanical point of view this instrument is a much simpler device. There are no lenses, and the mechanical parts are very simple (see Fig. 17). Essentially the colorimetric components consist of a set of glass colour filters, the source and the glass panel at the viewing aperture. In all models of the P-N anomaloscope three colour equations are utilized, and in theory all known colour defects can be tested on this instrument, not only in their dichromatic form but also at the anomalous stage. (See Fig. 12 showing loci for the equation both for P-N and the Nagel anomaloscopes.)

Current models utilize Chance-Pilkington glass filters to make up the primaries of the colour mixtures. Colorimetric measurements of these at the viewing aperture of the anomaloscope show that they are desaturated in relation to the values given by the makers for filters alone. Table 4 shows the C.I.E. coordinates and dominant wavelength (λ_D) as calculated from the position of illuminant A. The retinal illumination at modal match settings for



1 SOURCE 2 FILTERS 3 SHUTTERS

4 INTEGRATING CHAMBERS 5 ADJUSTABLE VIEWING APERTURE

FIG. 17. The P-N anomaloscope.

the three equations is around 70 and 80 trolands, and so is within the photopic range of adaptation. The viewing aperture can be varied from 40 minutes to $3\frac{1}{2}^\circ$ and viewing is done at one metre. Other subtense may be utilized, but the one most frequently used is the $1\frac{1}{2}^\circ$ measuring purely foveal vision.

Validation

Both anomaloscopes have proved to be valid measures of defective colour vision, and Willis and Farnsworth (1952) found the Nagel to be the most efficient of the five that they studied, that is the Nagel, Hecht-Shlaer, Bausch and Lomb, Double Dichroic Polaroid, and the Colour Temperature Meter Anomaloscopes. Green (1962) found complete agreement in classifying 21 red-green defectives between the Pickford-Nicolson and the Nagel

TABLE 4

C.I.E. COORDINATES AND λ_D (nm.) FOR THE PRIMARIES OF THE THREE COLOUR EQUATIONS (P-N ANOMALOSCOPE); MEASUREMENT BY VISUAL COLORIMETRY, ILLUMINANT A

Colour mixtures	C.I.E. coordinates								
	Primary 1			Primary 2			Standard		
	x	y	λ_D	x	y	λ_D	x	y	λ_D
Red + green = yellow	0.672	0.309	628.5	0.384	0.550	555.0	0.511	0.446	584.3
Yellow + blue = neutral	0.516	0.458	583.5	0.195	0.152	472.8	0.474	0.412	593.0
Green + blue = blue/green	0.379	0.548	552.5	0.172	0.131	473.3	0.302	0.400	493.5

anomaloscopes. In their original paper on the pilot model, Pickford and Lakowski (1960) state that the diagnosis obtained on the new model corresponded closely to results obtained on the older type of Pickford anomaloscope. Lakowski (1968d) also quotes a close correspondence between the anomaloscope findings and wavelength discrimination data obtained from the König-Helmholtz colour mixer; those subjects who had the best matches on the anomaloscope also had the best discrimination on the colour mixer and resembled those quoted by Wright (1946). The progressively poorer discrimination of some normal trichromats, typified by their enlarged matching ranges, was reflected in increased $\Delta\lambda$ at maxima of discrimination, that is, at blue, green, and red. There is also a close correspondence in their diagnosis of anomalous trichromats. When results were compared with those for the 100-Hue test, there was a similar correspondence between these diagnosed dichromats and the confusion axes. Other correlations between the performances of normal trichromats on the P-N and those on conventional colour vision tests have been described elsewhere by Lakowski (1965a, 1968d).

Norms for the P-N anomaloscopes have been established by Pickford and Lakowski (1960) and for the age variable by Lakowski (1968d) which in many respects are similar to those obtained from a larger sample tested on the older anomaloscope (Pickford, 1957; Lakowski, 1958, 1962). Norms for the Nagel are found in the works of Nelson (1938), Wright (1946), Rushton and Baker (1964), and for the age variable in the works of Obi (1954) and Boles-Carenini (1954).

Lakowski (1968d) states that quotients or *mid-matching points* are stable, showing no age variations around the most frequent ratio except among the young (particularly those under 15), where shifts in the red-green ratios, statistically significant at the 0.01 level, are obtained in relation to other populations. Boles-Carenini (1954), however, found that the means of the Rayleigh equations are displaced towards the green end in aged subjects. Lakowski (1962) also states that the magnitude of the standard deviations in the very young and in the older subjects is greater than is found in the 20-30 age group. As far as matching ranges (discrimination) are concerned, the red-green equation is the least affected by age deterioration, remaining stable until we come to the 55s and over, when the incidence of those with poor discrimination increases to 30%. On the other hand, matching ranges for yellow-blue and green-blue begin to increase as early as 30 years of age, where the percentage of subjects with poorer discrimination at, say, the age of 40 is already 30% higher than in the younger subjects, and reaches as high as 80% among the very old. Figure 18 gives

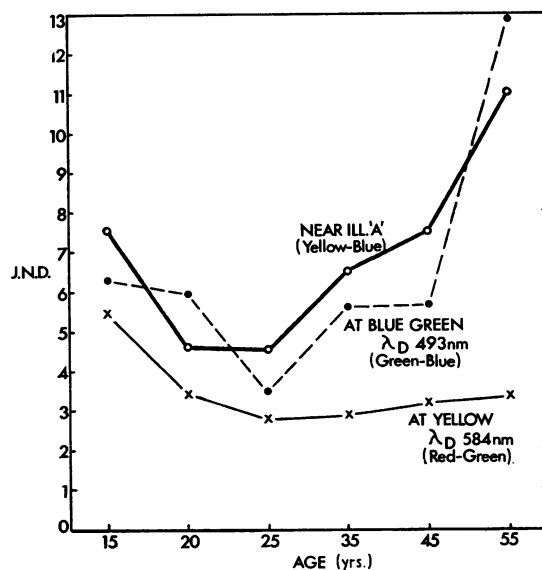


FIG. 18. Effect of age on discrimination (in j.n.d.s) — from a study using the P-N anomaloscope. Standards are indicated for each equation.

the mean matching ranges in j.n.d.s for three equations for the P-N anomaloscope for an age span from 15 to 55 years.

Of all the tests discussed here, the anomaloscopes are the most difficult to use, and training is necessary if they are to be used validly and efficiently. Only those who are seriously considering working in the field of colour vision should embark on using them. However, provided such skills are obtained, the anomaloscope has advantages which far outweigh any inconveniences in training. Colour vision, including the defects found among dichromats, anomalous trichromats, and the minor variations of normal subjects, can all be tested on a continuous scale. In addition, in anomaloscopes such as the P-N, red-green and yellow-blue defects and discrimination can be tested validly, and because a variable brightness adjustment is included for the standard yellow in both the Nagel and P-N anomaloscopes the classification of deutan and protan defects is categorical.

Special tests of colour aptitude

The requirements of industry, transport, and the professions have yet to be adequately stated, though in certain branches specific testing procedures and requirements for 'proper' colour efficiency are laid down. For example, the Giles-Archer and the Edridge-Green lanterns (Giles, 1960), are used for selecting air and naval personnel and for testing the ability to recognize small colour signals at long

distances. In some occupations, it is especially important that no dichromats, particularly of the red-green type, should be admitted. However, there are a vast number of occupations where this type of exclusion would not be adequate for selecting personnel, for, in industry especially, it seems more important to discover whether a man is fitted for a particular job than to classify him as either normal or colour defective. What is necessary in most situations is to establish whether he has the required skill to deal with a particular colour task. Such 'job analyses' have not yet been fully established, although the need for them has been stressed by writers such as Riddell *et al.* (1949), Colour Group (1946), Pickford (1955), Cavanagh (1955), Sloan and Altman (1951), and Lakowski (1968b).

Recently a number of tests have appeared which claim to test specific colour abilities such as colour aptitude and colour memory. Though they were designed only between 10 and 20 years ago these tests are becoming more widely accepted and their usefulness is now appreciated.

Burnham-Clark-Munsell colour memory test (B.C.M.T.)

To fill an important gap in an already existing battery of tests, Burnham and Clark (1955) devised their memory test, because they claimed that (apart from closely juxtaposed colour matches in small fields) all comparisons involved an element of memory. When a colour inspector compares a colour sample in one place with a colour standard elsewhere, he has to be able to recall his impression of the sample for some time if he is to compare it with the standard. The artist needs to remember the colours of the landscape while he paints, and the housewife frequently has to remember the colour of the walls when buying curtain material.

The B.C.M.T. is administered under standardized lighting conditions outlined elsewhere, and consists of a wheel on which the test chips or caps and the comparison chips are mounted. All this material is enclosed in a carrying case, the top surface of which serves as the immediate visual surround. Comparison chips consist of 43 of the odd number of hues of the Munsell series (as used in the 100-Hue test). Duplicates of 20 of these hues are used as test chips and duplicates of a further two hues are used as practice chips, all of which are mounted on two concentric circles on a freely rotating wheel.

The subject is presented with one surface hue for 5 seconds and after an interval of 5 seconds he is required to select from the range of hue chips one which resembles the first hue shown. This is repeated for as many as 20 of the hue chips and the subject is scored on his accuracy of memory choice. The test takes 15 to 20 minutes to administer individually. Scores obtained from a large population are normally

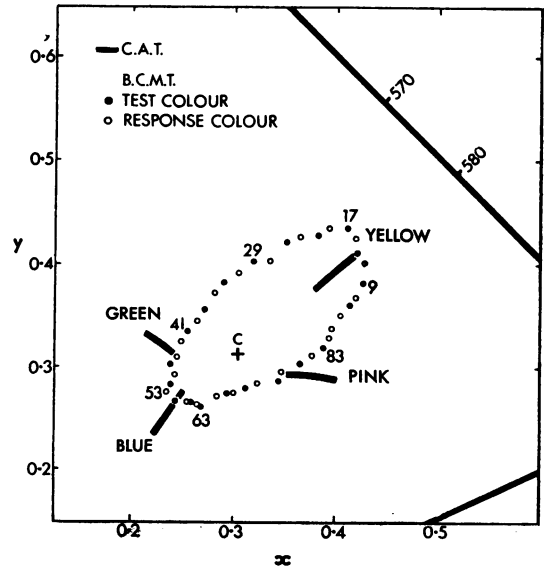


FIG. 19. Loci for the four C.A.T. series and chromaticities for the test and response colours of the B.C.M.T. (cap numbers as in the 100-Hue test).

distributed around a modal score of 20 to 24 according to the population studied, and it is possible to place a score within a percentile ranking order although in practice people are placed into three groups according to their hue-memory discrimination (a) superior, (b) normal, and (c) poor. The lower the score the better the performance. The B.C.M.T. is not a test for colour confusion and therefore should not be used to test for congenital colour defects. The chromaticities for the various chips in the C.I.E. diagram are quoted in Fig. 19 along with those for C.A.T.

Little has been done to validate this test against specific occupations so norms would have to be found empirically, but the information in Table 5 can be taken as a guide to the probable mean scores for certain populations. Data for industrial populations are from Burnham and Clark (1955), others were obtained from studies undertaken at the Visual Laboratory and these are arranged according to age to show the effect this variable has on the scores. It appears that the 'age effect' operates by showing an improvement in performance as age increases from 12 to 25 years. That a decline in performance takes place with oncoming old age has yet to be proven, but it remains a tenable hypothesis. It was also found that not all test items are equally difficult. Some chips are clearly easier to recall than others, and this difficulty is not randomly distributed but follows a distinct pattern. Consistent hue shifts, which appear to be essentially bipolar, are apparent in all populations so far studied, with one shift

TABLE 5
ERROR SCORES OBTAINED ON THE B.C.M.T.
BY VARIOUS POPULATIONS

N	Age	Population	B.C.M.T. score	
			Mean	SD
30	12-13	School children	41.5	14.7
106	13-16	Printers' apprentices	28.2	7.7
40	15-16	Secondary pupils	24.0	4.6
14	17-19	Students (defective R-G)	24.2	5.5
42	17-19	Students (normal vision)	23.5	4.5
27	17-22	Art college students	20.5	6.0
20	25	Industrial workers - minimal training in colour work	21.7	6.0
20	25	Industrial work - highly trained in colour work	19.4	5.7
130	25	Technical and clerical workers	20.0	—

toward the red end of the colour continuum and the other towards the green.

Though performance on the B.C.M.T. and tests such as the 100-Hue, C.A.T., and the P-N anomaloscope, are positively correlated, these relationships are small and not very significant and so it is fair to assume that the test measures an aspect of colour vision that the other tests do not. It can therefore be safely included in a battery of tests used to detect aspects of colour vision not measured by the more conventional tests.

Inter Society Colour Council colour aptitude test (I.S.C.C. - C.A.T.)

This test was designed by Dimmick (1946) to help select colour workers involved in matching and sorting. It was his contention that for such a test . . . 'some kind of matching technique was indicated, in which a degree of manipulation played a part'. His aims were to test innate and acquired colour skills, which can be developed through training, and essentially he is testing the ability to discriminate small differences in saturation at four points (or areas) in the colour space.

The test consists of some movable and some immovable coloured chips, 35 × 45 mm. in size, subtending an angle of about 4½° at the retina when viewed at a distance of 30 to 40 cm. The 48 immobile chips are fixed on to a panel in four horizontal rows, the first blue, then red, green, and yellow at the bottom. The different hue saturations are randomly arranged. The corresponding 48 movable chips are kept in a plastic chip dispenser (Dimmick, 1956). The colour task consists of matching the individual chips out of the dispenser with the corresponding ones on the board, where a perfect match scores

3 points, and scores of 2 and 1 points can also be gained for near matches. The full scores for the four hues are as follows: blue - 22, red - 29, green - 28, and yellow - 29 points, with a maximum of 109 points. Frequently scores are expressed as a percentage of the total possible within the range. The position of the four hue series are indicated in Fig. 19 (for additional colorimetric and spectrophotometric data, see Lakowski (1966), Golden Gate Society for Coatings Technology (1966)). It should be stated here that the colour differences involved in near matches are very small indeed, being only around 0.8 to 3 N.B.S. units.

For this test, illumination is critical. It should consist of a type 'C' illuminant giving initial illuminance of about 300 lux of 28 ft/candles. But good daylight of comparable intensity is an acceptable substitute. In the United States the Macbeth Daylight Lamp is recommended. Because the coloured chips have a glossy surface, viewing conditions must eliminate specular components and therefore C.I.E. recommended viewing should be employed, the panel placed at an angle of 45° to the illuminant and the subject's position maintained at 'normal'.

The test is rather difficult and time consuming and it is always good practice to warn the subject of this. It takes about 40 minutes to complete although some subjects take one hour or more but there is no correlation between the time taken to finish the test and the final score.

The C.A.T. has been used more extensively than the memory test and certain norms have been collected. Some users have mistakenly expected this test to detect colour defectives but the test was not designed to detect colour confusion, although some of its elements align themselves along certain confusion lines for each type of dichromat (Lakowski, 1966). Extensive empirical studies to find out if the test can pinpoint colour vision defects were undertaken on a few dichromats by Tilleard (1958), Adams (1958), and Green (1962), and on a larger number of colour defectives by Walraven, Leebeck and Bouman (1956) and Verriest, Seki, Metsala, and Buysens (1964). All these investigators agree that it would be wrong to administer the C.A.T. without first screening the subject for colour deficiency, since the total score was not a reliable indication of defect, although they found that the red and green scores of defectives were lower than those of normals.

Data from a study by Gilbert (1957) on 160 males and 195 females, arranged in decades from 10 to 90 years of age, tell us something about the effect of age and sex on C.A.T. scores. Her main conclusions were that scores on the C.A.T. reach a peak in the mid-twenties and then gradually diminish, and that although the decrease is consistent, it is so gradual that the differences between adjacent decades are not

TABLE 6
MEAN SCORES OBTAINED ON THE C.A.T., BY VARIOUS POPULATIONS

<i>N</i>	<i>Mean score</i>	<i>Age and type of subject</i>	<i>Investigator</i>
106	51.1	13-16, printers' apprentices	Lakowski (1969)
355	53.6	Age study, 10-90 years, no colour experience	Gilbert (1957)
25	59.1	16-52, average 28 years, colour workers	Lakowski (1968c)
48	59.2	Adult (?) untrained subjects	Walraven <i>et al.</i> (1956)
112	60.9	14-53, average 28 years, random population	Verriest <i>et al.</i> (1964)
147	66	16-60 years, students of architecture and colour workers	Little (1967)
200	70	Adult colour workers	Adams (1958)
200	70	Adult (?) varied	Dimmick (1954)
698	70	Adult (?) concerned with colour problems	Dimmick (1956)
210	72.7	Adult colour trained	Tilleard (1958)

significant. At all ages in this population discrimination was better in yellow and red than in green and blue. Females had higher total scores, but this was not true for all colours or for all age decades.

Several studies of colour normals were made using the C.A.T., and mean scores were found to vary from 53 to 72, a rather large scatter. This may be attributable to small colorimetric differences which possibly exist between individual sets of the C.A.T. A more probable cause is that the test was not administered under identical lighting conditions nor in an identical manner. Table 6 shows the results of a number of investigations giving mean scores, size of sample, age, and a brief description of the type of subject used.

The validity of the C.A.T. for testing for specific occupations is not yet known and still has to be established empirically, but correlations between the scores obtained on this test and on the B.C.M.T., 100-Hue, and P-N anomaloscope indicate that there is no relationship between C.A.T. results and the others, and therefore it can again be assumed, as in the case of the colour memory, that this test measures some ability that the others do not. A four-point classification of total individual scores can be used and two such have been suggested by Dimmick (1954) and by Hess (1957).

The C.A.T. test is used extensively for selecting colour workers in the U.S.A., and it has recently been suggested (I.S.C.C. Newsletter No. 173) that if

selection is made on the basis of C.A.T. scores, only those with a score of 75 or over should be chosen to become colour matchers.

Discussion

No direction has yet been given on what tests to use. This was deliberate. The reader was presented with the relevant facts – how each test was designed, how it detects defects, and why it is able to measure individual differences in colour discrimination. It was pointed out that the tests are not perfect; that it is sometimes difficult to achieve what is expected; and that this may depend on variables not involving colour. It is hoped that those already using such tests will understand them a little better, that their results will be more reliable, and that they will eliminate any wrong inferences they may have habitually been making. Newcomers should find their work made easier. As far as the actual choice of tests is concerned there is obviously no perfect selection. For example, it might be thought best to use the most sophisticated tests (those which measure a variety of colour attributes) but high costs, length of testing time, and the skill required in applying them all might prevent them from being chosen. The following discussion is intended to clarify the factors involved in such a choice.

Colour vision requirements

This requires careful analysis, and the habit of looking upon colour vision problems as simple and requiring little attention should be deplored. A kind of job analysis is always called for, and where this is not readily available an attempt should be made from a common sense point of view. First, it should be decided whether to eliminate only those with major defects or whether those with the best discrimination are wanted. This distinction is important as the exclusion of dichromats does not ensure that those with good colour discrimination will be selected. It might be that after more careful con-

TABLE 7
FOUR-PART CLASSIFICATION OF SCORES FOR C.A.T.

<i>Investigator</i>	<i>Four-point scale</i>			
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
Dimmick	85-95 Excellent	75-85 Good	65-75 Satisfactory	55-65 Doubtful
Hess	84-109	72-83	49-71	0-48

sideration it will become obvious that not all defectives need be excluded, or alternatively that it is unnecessary to select only the best for the task. There are professions and skills in which it is apparently an advantage to have good discrimination although it is a mistake to think that a defect would be an insurmountable handicap. Take, for example, professions like architecture, metallurgy, and crystallography, or the biological sciences, where staining techniques are used, and analytical chemistry in which titrations are involved. In cases like these, colour vision defects, especially dichromatism, are a handicap since subjects who are unaware of their defects can be blamed for lack of attention, stupidity or ignorance if they mistake one colour for another. It would probably be unfair to eliminate colour defectives from such occupations, but it is always good practice to let people know what type of discrimination they have in order to prevent such unpleasant occurrences, and where the eye is to be used as a sort of 'null instrument' they must either ask others to help, pay particular attention to the task, or try to reduce as far as possible the dependence on colour clues for solutions.

Of course there are situations where only those with the best discrimination are wanted and where this is so we must make absolutely certain that only those with such ability are selected.

Available tests

Even when colour vision requirements are known we still have to consider what tests are available and the amount of money and time at our disposal. Perhaps it should be stressed that a prevalent attitude, which assumes that any test will do and that the cheaper it is and the less time it takes the better, is not conducive to an adequate analysis of colour vision. This is clearly the wrong approach. Of course if we only wish to eliminate red-green dichromats then pseudo-isochromatic plates costing about £6 and administered in a few minutes are sufficient, though it should be realized that some of those who are outright failures on these tests will have only minor defects. In order to avoid this, greater expenditure is needed and longer time must be allowed for administration. We must also consider whether the tester possesses or is willing to acquire the skills which would enable him to administer the more complicated tests validly. Again it is the experience of the writer that most of those involved in colour vision assessment look on it as a sort of Cinderella that needs scant attention and no skill whatsoever. It must be stressed that this is not so; that if procedure is standardized and a few techniques learned much more will be achieved than the mere knowledge of whether the person is colour defective or not. Such competence gives a degree of certainty that would not have been possible before.

Thus tests of hue discrimination or colour aptitude and the anomaloscopes in particular should not be administered by first reading the manual. It is the practice at the Visual Laboratory when training people to administer these tests that a number of subjects, usually at least 20, must be tested in a variety of tests before the experimenter is regarded as having some competence in handling the material. If this skill is not acquired, two things can happen. On the one hand, the tester will be clumsy and awkward in dealing with his subjects, many of whom will be very sensitive about being found to be defective. On the other hand, if the more advanced tests are used, it will be found that wrong or inadequate information will be obtained as a result of the test situation being made too easy or too difficult. Since the results of such administration will not be comparable with the norms obtained by others, it is strongly advised that every examiner should spend some time acquiring the techniques of administration until he acquires some finesse.

Actual testing procedure

Let us begin by saying that the four PIC tests discussed in this paper are capable of detecting red-green dichromats and if administered in northern day-light they will be the cheapest and easiest method of testing. If, on the other hand, an artificial standard illuminant 'C' is to be used then an additional £20 to £40 will have to be expended. If we must also eliminate the anomalous, those with tritan defects, or those with possible acquired defects, then the simple PIC tests would not be adequate. In cases where anomalous trichromats have to be detected the use of hue discrimination or the specialized tests would add nothing to our diagnosis as they are not sufficiently sensitive. An anomaloscope which can distinguish between the various types and degrees of defect would be required. This immediately poses problems – one of cost since the two anomaloscopes described here would be priced between £150 and £350, and another of skill since the tester must have this and must be prepared to spend a certain time, say a few weeks, acquiring it.

When selecting people with good discrimination two things have to be considered. We have already seen that specialized tests cannot distinguish properly between congenital colour defectives and those with none, so we must first establish that in the population to be tested for specific abilities all such defectives have been eliminated by other means. For this we would require a pseudo-isochromatic test to begin with, followed by the specialized tests of colour discrimination, colour memory or aptitude, and we would have to know exactly what colour ability we were looking for. That this is an important problem can be seen from a recent study by

Lakowski (1969) where the effect of psychological variables and the inter-correlation between colour vision tests was assessed for normal trichromats. It was found from the results of factor analysis that each test was measuring attributes specific to itself and from this it is inferred that no one test can be regarded as a substitute for another. Let us enlarge on this. In the first place, as far as the psychological factors of personality and intelligence are concerned, though there was a significant relationship between some colour vision tests and cognitive variables and others with personality variables, it was found that actual variance for a given colour test was produced by something specific to that test and that variance due to other factors was small, accounting for about 4-16% of the total. For example, in the 100-Hue test, 84% of the variance was produced by something specific to this test, and the remainder, that is 16%, was made up of different factors; 7½% was accounted for by what this test has in common with other colour vision tests while 4½% and 4% respectively were explained by what it has in common with cognitive and personality factors. It was also shown that variance for cognitive and personality factors is about equal for the 100-Hue and the B.C.M.T. test while the cognitive loading is much smaller for the C.A.T. and the P-N anomaloscope which, in turn, are more loaded on personality aspects. However, from the study it is obvious that these extraneous variances are relatively insignificant compared with the total for each test. This is quite an important finding as far as our problem is concerned for if we want to select people with the best possible ability we have to know exactly which ability we are interested in and so we come back to the importance of job analysis. Remembering that each test measures a specific ability we have to be clear in our minds what we want from our colour worker or colour matcher. If the main element in the task is a good memory for colours, then obviously the B.C.M.T. test is the one required. If, on the other hand, the task involves the perception of small brightness or saturation differences, for example as found with machine-press operators, then the test we want to apply is the C.A.T.; and, lastly, where a premium is put on hue discrimination and where small differences must be distinguished, then the 100-Hue test should be employed.

We have seen that anomaloscopes detect colour confusion and can measure discrimination in certain parts of the colour space, but we should also remember that this test is most suited where metamerism plays an important role. Thus the instrument can be used as an overall test of colour vision and colour discrimination where the information obtained will be most relevant to colour mixing and to matching of metameric pairs as found in the textile and printing industries.

Perhaps the effect of ageing and acquired dyschromatopsias should again be emphasized since this calls for re-testing at certain ages. For example, a man of 40 might be rather old to be made a foreman in a colour or dyestuff workshop and if he is chosen then he should be re-tested at short intervals, say, of about five years. It should also be remembered that if industrial workers either have or had any eye diseases or have complications such as diabetes or arthritis there is a possibility that their discrimination has deteriorated. It must be stressed that pseudo-isochromatic tests are both inadequate and inappropriate for detecting such changes and therefore the emphasis is again on colour discrimination tests or anomaloscopes.

Battery of tests

It will be plain by now that the ideal is to have a set of colour vision tests so that each situation can be met. The question of cost must be borne in mind. If it is only necessary to exclude major defectives and then keep the cost as low as possible, it is perhaps advisable to use two sets of pseudo-isochromatic plates – one of the modern types such as the T.M.C. or AO H-R-R and the other from among the Ishihara or Dvorine tests, or the Stilling tables. A more sophisticated and useful combination would be one pseudo-isochromatic test plus a colour discrimination test like the 100-Hue. Here, with a little skill, all variations in colour vision could be detected for although it has been said that from 100-Hue results alone we will have no indication of simple anomalous vision, yet, used in conjunction with the plates, it will be found that with a little practice a fairly certain diagnosis can be made. The third combination would, of course, involve very specialized tests and it should be remembered that even here we would still have to administer a pseudo-isochromatic or other test for screening purposes. It would therefore be necessary to have that test plus any of the specific ones already mentioned. In large establishments which undertake testing for different industries or occupations or where an industrial consultant is employed, the ideal solution would be to have a battery of tests which include an anomaloscope as well as a colour discrimination test and some PIC tests. But, of course, in the last analysis all these tests will only be as effective as the astuteness of the user can make them. Though papers like this are helpful, there is no substitute for practice.

I wish to thank Miss C. Liston for secretarial assistance and Mr. J. Rae for preparing the diagrams.

References

- Adams, J. M. (1958). Some results obtained with the I.S.C.C. colour aptitude test. *J. Oil Col. Chem. Ass.*, 41, 807-813.

- Belcher, S. J., Greenshields, K. W., and Wright, W. D. (1958). Colour vision survey using the Ishihara, Dvorine, Boström and Kugelberg, Boström, and American-Optical Hardy-Rand-Rittler tests. *Brit. J. Ophthalmol.*, **42**, 355-359.
- Boles-Carenini, B. (1954). Del comportamento del senso cromatico in relazione all'età. *Ann. Otol.*, **80**, 451-458.
- Brown, P. K., and Wald, G. (1964). Visual pigments in single rods and cones of the human retina. *Science*, **144**, 45-52.
- Burnham, R. W., and Clark, J. R. (1955). A test of hue memory. *J. appl. Psychol.*, **39**, 164-172.
- , Janes, R. M., and Bartleson, C. J. (1963). *Color - A Guide to Basic Facts and Concepts*. Wiley, New York.
- Cavanagh, P. (1955). The Ishihara test and defects of colour vision. *Occup. Psychol.*, **29**, 43-57.
- Colour Group (1946). *Report on Defective Colour Vision*. Physical Society, London.
- Cox, J. (1961a). Colour vision defects acquired in diseases of the eye. *Brit. J. physiol. Opt.*, **17**, 195-216; **18**, 3-32; and **18**, 67-84.
- (1961b). Unilateral color deficiency, congenital and acquired. *J. opt. Soc. Amer.*, **51**, 992-999.
- Crawford, B. H., Ed. (1968). *Techniques of Photostimulation in Biology*. North Holland Publ. Co.
- De Valois, R. L. (1960). Color vision mechanisms in the monkey. *J. gen. Physiol.*, **43**, July Suppl. pp. 115-128.
- , Jacobs, G. H., and Jones, A. E. (1963). Responses of single cells in primate red-green color vision system. *Optik (Stuttgart)*, **20**, 87-98.
- Dimmick, F. L. (1946). Color Aptitude Test. *J. appl. Psychol.*, **30**, no. 1, pp. 10-22.
- (1954). Factors in the application of the Colour Aptitude Test. *Off. Digest.*, pp. 1265-1270.
- (1956). Specifications and calibration of the 1953 edition of the Inter-Society Colour Council Aptitude Test. *J. opt. Soc. Amer.*, **46**, 389-393.
- Farnsworth, D. (1943). The Farnsworth-Munsell 100-hue and dichotomous test for color vision. *Ibid.*, **33**, 568-578.
- (1955). Methodological requirements in the study of normal and defective color vision. In *Recent Developments in Color*. (Proc. 14th int. Congr. Psychol., Montreal 1954). *Acta Psychol.*, **11**, 223.
- (1957). *The Farnsworth-Munsell 100-Hue Test Manual*. Munsell Color Co., Baltimore, Maryland, U.S.A.
- François, J., and Verriest, G. (1957). Les dyschromatopsies acquises. *Ann. Oculist. (Paris)*, **190**, 713-746, 812-859, and 893-943.
- , — (1961). On acquired deficiency of colour vision. *Vision Res.*, **1**, 201-219.
- Galifret, Y., Ed. (1960). *Mechanisms of Colour Discrimination*. Pergamon Press, Oxford.
- Gilbert, J. G. (1957). Age changes and color matching. *J. Geront.*, **12**, 210-215.
- Giles, G. H. (1960). *The Principles and Practice of Refraction*. Hammond, Hammond, London.
- Golden Gate Society for Coatings Technology (1966). Color Instruments take the visual Color Aptitude Test. *J. Paint Technology*, **38**, 564-574.
- Graham, C. H., Ed. (1965). *Vision and Visual Perception*. Wiley, New York.
- Granit, R. (1947). *Sensory Mechanisms of the Retina*. Oxford University Press, London.
- (1955). *Receptors and Sensory Perception*. Yale University Press, New Haven, Conn.
- Green, M. (1962). Results of testing a group of colour defective subjects with several tests of colour vision and colour aptitude. *Farbe*, **11**, 87-92.
- Grossman, S. P. (1967). *A Textbook of Physiological Psychology*. Wiley, New York.
- Grütznher, P. (1962). Erworbene Farbensinnstörungen bei Netzhautdegenerationen. *Farbe*, **11**, 35-44.
- Guild, J. (1925-26). A trichromatic colorimeter suitable for standardization work. *Trans. opt. Soc. (London)*, **27**, 106-129.
- (1931). The colorimetric properties of the spectrum. *Phil. Trans. roy. Soc. (London) A*, **230**, 149-187.
- Hardy, L. H., Rand, G., and Rittler, M. C. (1945). Tests for the detection and analysis of color blindness. I. The Ishihara test: an evaluation. *J. opt. Soc. Amer.*, **35**, 268-275.
- Helmholtz, H. von (1892). Versuch das psychophysische Gesetz auf die Farbenunterschiede trichromatischer Augen anzuwenden. *Z. Psychol. Physiol. Sinnesorg.*, **3**, 1-20.
- Hering, E. (1878). *Zur Lehre vom Lichtsinne*. Gerolds, Vienna.
- Hering, E. (1920). *Grundzüge der Lehre vom Lichtsinn*. Julius Springer, Berlin.
- Hess, M. (1957). Personal experience with the I.S.C.C. colour aptitude test. *J. Oil Col. Chem. Assoc.*, **40**, 136-142.
- Hurvich, L. M., and Jameson, D. (1955). Some quantitative aspects of an opponent-colors theory, II. Brightness, saturation and hue in normal and dichromatic vision. *J. opt. Soc. Amer.*, **46**, 602-616.
- , — (1960). Color vision. In *Ann. Rev. Psychol.*, **11**, 99-130. Inter Society Color Council. News letter No. 173.
- Jaeger, W. (1956). Defective colour-vision caused by eye disease. *Trans. ophthalm. Soc. U.K.*, **76**, 477-489.
- Judd, D. B. (1935). A Maxwell triangle yielding uniform chromaticity scales. *J. opt. Soc. Amer.*, **25**, 24-35.
- (1944). Standard response functions for protanopic and deuteranopic vision. *J. Res. nat. Bur. Standards*, **33**, 407; RP 1618.
- (1949). Current views on colour blindness. *Docum. Ophthalm. (den Haag)*, **3**, 251-288.
- (1958). Basic correlates of the visual stimulus. In *Handbook of Experimental Psychology*, S. S. Stevens (Ed.), pp. 811-867. Wiley, New York.
- (1961). Blue-glass filters to approximate the blackbody at 6500°K. *Farbe*, **10**, 31-36.
- (1966). Fundamental studies of color vision from 1860 to 1960. *Proc. Nat. Acad. Sci. (Wash.)*, **50**, 1313-1330.
- , and Wyszecki, G. (1963). *Colour in Business, Science and Industry*, 2nd ed. Wiley, New York.
- Kalmus, H. (1955). The familial distribution of congenital tritanopia. *Ann. Hum. Genetics*, **20**, 39-56.
- (1965). *Diagnosis and Genetics of Defective Colour Vision*. Pergamon Press, Oxford and London.
- Kherumian, R., and Pickford, R. W. (1959). *Hérédité et Fréquence des Anomalies Chromatiques du Sens Chromatique*. Vigot Frères, Paris.
- Kinnear, P. R. (1965). The colour discrimination of diabetics. Unpublished M.Sc. thesis, Edinburgh University.
- Kries, J. von. (1905). Die Gesichtsempfindungen. In *Handbuch der Physiologie des Menschen*, ed. Nagel, W., Vol. 3, 109-281, Vieweg, Braunschweig.
- (1924). In *Treatise on Physiological Optics*. H. L. F. von Helmholtz. (J.P.C. Southall trans.). Optical Society of America, Rochester, New York State.
- Koellner, H. (1912). *Die Störungen des Farbensinnes. Ihre klinische Bedeutung und ihre Diagnose*. Karger, Berlin.
- König, A. (1903). *Gesammelte Abhandlungen*. Barth, Leipzig.
- Lakowski, R. (1958). Age and colour vision. *Advanc. Sci.*, **15**, 231-236.
- (1962). Is the deterioration of colour discrimination with age due to lens or retinal changes? *Farbe*, **11**, 69-86.
- (1964). Age and colour vision. Unpublished Ph.D. thesis, Edinburgh University.
- (1965a). Testing of colour vision in prospective printers' apprentices and the problems this presents in selection. *Brit. J. Physiol. Opt.*, **22**, 10-32.
- (1965b). Colorimetric and photometric data for the 10th edition of the Ishihara plates. *Ibid.*, **22**, 195-207.
- (1966). A critical valuation of colour vision tests. *Ibid.*, **23**, 186-209.
- (1968a). The Farnsworth-Munsell 100-hue test. *Ophthalm. Optician*, **8**, 862-872.
- (1968b). Colour matching ability: can it be measured? *J. Soc. Dyers Col.*, **84**, 3-9.
- (1968c). Small colour vision variations and their effect in visual colorimetry. In *Colour Measurement in Industry*, pp. 167-194. The Colour Group, London.
- (1968d). The Pickford-Nicolson Anomaloscope - Calibration validation and population norms. Second Scottish Symposium on Colour, Scottish Section of the Colour Group. 5 and 6 September. *Brit. J. Physiol. Opt.* In press.
- (1969). Psychological variables in colour vision testing. Proceedings of the 1st International Colour Congress, Stockholm, 1969. M. Richter, Masterschmidt, Verlag, Göttingen. In press.
- Little, A. (1967). The measurement of color discrimination ability. *Color Engineering*, **5**, 48-53.
- Le Grand, Y. (1968). *Light, Colour and Vision*, 2nd ed. Chapman and Hall, London.
- Marks, W. B., Dobbelle, W. H., and MacNichol, E. F. (1964). Visual pigments of single primate cones. *Science*, **143**, 1181-1183.
- Maxwell, J. C. (1890). On the theory of colours in relation to colour-blindness. In *The Scientific Papers of James Clerk Maxwell*, vol. 1, pp. 119-125, ed. Niven, W. D. Cambridge University Press, London.
- Müller, G. E. (1924). *Darstellung und Erklärung der verschiedenen Typen der Farbenblindheit*. Vandenhoeck and Ruprecht, Göttingen.
- (1930). Ueber die Farbenempfindungen, Bd. 1 and 2. *Z. Psychol. Physiol. Sinnesorg.* Ergänzungsbd., **17**, 1-430; **18**, 435-647.
- Nagel, W. A. (1907). Zwei Apparate für die augenärztliche Funktionsprüfung. Adaptometer und kleines Spektrophotometer (Anomaloskop). *Z. Augenheilk.*, **17**, 201-222.

- Nelson, J. H. (1938). Anomalous trichromatism and its relation to normal trichromatism. *Proc. phys. soc. Lond.*, **50**, 661-690.
- Nickerson, D., and Stultz, K. F. (1944). Color tolerance specification. *J. opt. Soc. Amer.*, **34**, 550-570.
- Obi, S. (1954). Study on the color sense of the aged. *Acta Soc. Ophthal. Jap.*, **58**, 451-454. (In Japanese).
- Ohta, Y. (1966). The improvement of 40 Hue test. *Ibid.*, **70**, 798-801. (In Japanese, English summary.)
- Ostwald, W. (1931-33). *Colour Science: A Handbook for Advanced Students*. Winsor and Newton, London.
- Peters, G. A. (1954). The new Dvorine color perception test. *Optometric Weekly*, No. 11, pp. 1801-1803.
- Pickford, R. W. (1949). Total colour blindness of hysterical origin. *Brit. J. med. Psychol.*, **22**, 122-128.
- (1950). A study of the Ishihara test for colour blindness. *Brit. J. Psychol.*, **40**, 71-80.
- (1951). *Individual Differences in Colour Vision*. Routledge and Keegan Paul, London.
- (1955). Weak and anomalous colour vision in industry and the need for adequate tests. *Occup. Psychol.*, **29**, 182-192.
- (1957). A practical anomaloscope for testing colour vision and colour blindness. *Brit. J. physiol. Opt.*, **14**, 2-26.
- (1958). A review of some problems of colour vision and colour blindness. *Advanc. Sci.*, **15**, 104-117.
- , and Lakowski, R. (1960). The Pickford-Nicolson anomaloscope. *Brit. J. Physiol. Opt.*, **17**, 131-150.
- Pitt, F. H. G. (1935). Characteristics of dichromatic vision. *Med. Res. Coun. spec. Rep. Ser.* No. 200.
- Rayleigh, Lord (1881). Experiments on colour. *Nature (Lond.)*, **25**, 64-66.
- Riddell, W. J. B. et al. (1949). Discussion on colour vision in industry. *Proc. roy. Soc. Med.*, **43**, 145-150.
- Ripps, H., and Weale, R. A. (1963). Cone pigments in the normal human fovea. *Vision Res.*, **3**, 531-543.
- Roth, A. (1966). Le test 28 hue selon Farnsworth. *Bull. Soc. Ophthal. Fr.*, **66**, 231-238.
- Ruddock, K. H. (1965). The effect of age upon colour vision. *Vision Res.*, **5**, 37-45; 47-58.
- Rushton, W. A. H. (1958). Visual pigments in the colour blind. *Nature (Lond.)*, **182**, 690-692.
- (1962). Visual pigments in man. *Sci. Amer.*, **207**, no. 5 (Nov.), pp. 120-132.
- , and Baker, H. D. (1964). Red-green sensitivity in normal vision. *Vision Res.*, **4**, 75-85.
- Said, F. S., and Weale, R. A. (1959). The variation with age of the spectral transmissivity of the living human crystalline lens. *Gerontologia (Basel)*, **3**, 213-231.
- Sidowski, J. B., Ed. (1966). *Experimental Methods and Instrumentation in Psychology*. McGraw-Hill, New York.
- Sloan, L. L., and Altman, A. (1951). *Evaluation of H-R-R plates for Measuring Degree of Red-Green Deficiency*. Wilmer Ophthal. Instit. Tech. Report-Project N 60 NR 234-07.
- Smith, H. C. (1943). Age differences in color discrimination. *J. gen. Psychol.*, **29**, 191-226.
- Stiles, W. S. (1952). Colour vision: a retrospect. *Endeavour*, **11**, 33-40.
- , and Crawford, B. H. (1933). The luminous efficiency of rays entering the eye at different points. *Proc. roy. Soc.*, **112B**, 428-450.
- Svaetichin, G., and MacNichol, E. F. (1958). Retinal mechanisms for chromatic and achromatic vision. *Ann. N.Y. Acad. Sci.*, **74**, 385-404.
- Tiffin, J., and Kuhn, N. S. (1942). Color discrimination in industry. *Arch. Ophthal. (Chic.)*, **28**, 851-859.
- Tilleard, D. L. (1958). Colour Aptitude Test: analysis of scores. *J. Oil Col. Chem. Ass.*, **41**, 797-806.
- Trendelenburg, W. (1929). Zur Diagnostik des abnormalen Farbensinnes. *Klin. Mbl. Augenheilk.*, **83**, 721-732.
- Umazume, K., and Harutake, M. (1962). 'Tokyo Medical College Color Vision Test'. *Farbe*, **11**, 45-47.
- Verriest, G. (1963). Further studies on acquired deficiency of colour discrimination. *J. opt. Soc. Amer.*, **53**, 185-195.
- (1964). Les déficiences acquises de la discrimination chromatique. *Mém. Acad. roy. Méd. Belg.*, ser. 2, 4, fasc. 5, pp. 35-327.
- (1968). Étude comparative des efficacités de quelques tests pour la reconnaissance des anomalies de la vision des couleurs. *Arch. Mal. Prof.*, **29**, 293-314.
- , Seki, R., Metsala, P., and Buysens, A. (1964). Sur le color aptitude test et sa valeur en tant qu'épreuve de détection des déficiences de la discrimination chromatique. *Biol. Jaarb. (Dodonaea)*, **32**, 237-242.
- , Vandevyvere, R., et Vanderdonck, R. (1962). Nouvelles recherches se rapportant à l'influence du sexe et de l'âge sur la discrimination chromatique, ainsi qu'à la signification pratique des résultats du test 100 hue de Farnsworth-Munsell. *Rev. Opt.*, **41**, 499-509.
- Wald, G. (1966). Defective color vision and its inheritance. *Proc. nat. Acad. Sci. (Wash.)*, **55**, 1347-1363.
- Walraven, P. L. (1966). A zone theory of colour vision. In *Proc. int. Col. Meeting*. Vol. 1, pp. 137-140. Ed. M. Richter, Muster-schmidt-Verlag, Göttingen.
- , Leebeck, H. J., and Bouman, M. A. (1956). *The ISCC Color Aptitude Test - the Interpretation of Some Testing Results*. Instituut voor Zintuigfysiologie. R.V.O.-T.N.O. Report No. WW-10.
- Weale, R. A. (1963). *The Aging Eye*. Lewis, London.
- Willis, M. P., and Farnsworth, D. (1952). *Comparative Evaluation of Anomaloscopes*. U.S. Navy med. Res Lab., Washington. Report 190.
- Wright, W. D. (1928-29). A re-determination of the trichromatic coefficients of the spectral colours. *Trans. Opt. Soc. (Lond.)*, **30**, 141-164.
- (1930). A re-determination of the mixture curves of the spectrum. *Ibid.*, **31**, 201-218.
- (1946). *Researches on Normal and Defective Colour Vision*. Henry Kimpton, London.
- (1952). The characteristics of tritanopia. *J. opt. Soc. Amer.*, **42**, 509-521.
- (1964) *The Measurement of Colour*. 3rd ed. Hilger, London.
- Wyszecki, G., and Stiles, W. S. (1967). *Color Science: Concepts and Methods, Quantitative Data and Formules*. Wiley, New York.
- Young, T. (1807). On the theory of light and colours. In *A Course of Lectures on Natural Philosophy and the Mechanical Arts*, Vol. 2, pp. 613-632. Johnson, London.
- Zanen, J. (1953). Introduction à l'étude des dyschromatopsies rétinienes centrales acquises. *Bull. Soc. belge Ophthal.*, no. 103, pp. 3-148.

Reference list of tests

1. AO H-R-R Pseudoisochromatic plates. (2nd ed. 1957). Hardy, L. H., Rand, G., and Rittler, M. C. American Optical Instrument Division.
2. Burnham-Clark-Munsell Color Memory Test. Burnham, R. W., and Clark, J. R. Munsell Color Co. Inc., Baltimore, 1, Maryland 21218; and Tintometer Sales Limited, Waterloo Road, Salisbury, England.
3. Color Aptitude Test (1953). Inter-Society Color Council. Federation of Paint and Varnish Production Clubs, Philadelphia 7, Pennsylvania.
4. Dvorine Pseudo-Ischromatic Plates (2nd ed. 1953). Dvorine, I. Waverley Press Inc., Baltimore, Maryland, U.S.A.
5. Farnsworth-Munsell 100-Hue Test. Farnsworth, D. Munsell Colour Co. Inc., Baltimore 1, Maryland 21218; Tintometer Sales Limited, Waterloo Road, Salisbury, England.
6. Ishihara Tests for Colour-Blindness. (10th ed.) Ishihara, S. S. Kanchara, Nippon Isho Shuppan Co., Japan.
7. Pickford-Nicolson Anomaloscope. Pickford, R. W. Rayner of Rayner and Keeler Ltd., London.
8. Tokyo Medical College Color Vision Test (1957). (1st ed.) Ophthalmology Dept., Tokyo Medical College. Murakami Color Research Laboratory, Tokyo, Japan.

Received for publication November 16, 1968.