
Theory and practice of colour vision testing: A review Part 1

R. LAKOWSKI

Visual Laboratory, Department of Psychology, Edinburgh University

Lakowski, R. (1969). *Brit. J. industr. Med.*, 26, 173-189. **Theory and practice of colour vision testing: A review Part I.** It is the concern of this paper to examine not only the effectiveness of tests for detecting colour confusion but also their usefulness in assessing colour vision generally. In part 1, problems of administration and age and the question of the basic elements of such tests are discussed. The existing theory of colour vision and colour defect is outlined and the principle of objective colour specification, which is believed to be valuable for understanding these tests, is introduced.

There are many tests for detecting colour defects and as their number increases it becomes progressively more important to find which are the most useful, valid, and reliable. This paper is an account of the more popular and more frequently used tests but it departs from the usual practice of dealing with colour vision from the point of view of colour defects alone. The main objective is to assess colour vision, and defects or the absence of good colour discrimination are regarded only as part of this larger problem. This approach is considered more useful and informative. It is more useful because the problem in industry is not one of excluding only colour defectives but also one of selecting those with either good or the very best possible discrimination, and from this point of view our approach puts this in its proper perspective. It is more informative because it deals with the whole problem of colour discrimination and its variations. With it we are better able to understand what can and cannot be done according to the type of visual physiology that we have inherited. And apart from dealing with colour, in any real situation, factors such as age, and the fact that a given task may involve cognitive problems outside the realm of colour vision, are all seen to play their part.

Both theory and practice are taken into account

because it is felt that such a complementary approach should in the long run yield better results. It is simply not enough to know how to administer and score a test. It is useful to know why one test is good and another less so, or why one is more appropriate than another for a certain selection procedure; and knowing why a test is effective increases our confidence in our work. It is here that the knowledge of the past 50 years helps to bridge the gaps in our experience and permits us to reach a higher degree of competence in administering these tests.

There are other reasons for presenting the subject-matter in a novel way. The first is based on the firm conviction that an individual's test performance can only be fully assessed when it is compared with results of large population studies. This is essentially a psychologist's approach, and in this sense we follow the tradition of Pickford who looked upon defects simply as part of the total problem of individual differences in colour vision. The second is that understanding the structure of a test, especially from the information provided by the science of colorimetry, is vital, since this is the best possible way of assessing a test objectively. In a sense, this is the approach of the physicist who always stresses the importance of the stimulus. It is hoped that by using both approaches in this paper we will achieve

a better and more realistic understanding of colour vision testing.

There are many ways of assessing colour vision. The research worker does so in one way, the consultant in another. Thus, from the research point of view, the efficiency of a subject's colour vision can be found by measuring his wavelength discrimination ($\Delta\lambda$) and assessing his relative luminosity (or visibility) function (V_λ) since these two together will indicate what basic type of visual system he possesses. For such work we require optically sophisticated instruments to provide pure spectral stimuli and this requires laboratory conditions and rather lengthy administration. It might be argued that such a sophisticated approach is wasteful when all we are interested in is diagnosing extreme deficiencies and that less sophisticated methods can therefore be used, and it is a fact that from wavelength discrimination studies it is almost impossible to distinguish deutan from protan defects. Yet a good anomaloscope can give results which, though unlike those described above, are nonetheless more useful for classifying a defect correctly. There are even simpler methods using a test involving surface colour where a reasonable assessment can be made. As this paper is not intended for research purposes but designed to be of practical value to those who have to select colour workers, the techniques we describe are mostly those used in this country, in Europe or in America by people involved in testing for vocational purposes. The actual choice depends on which tests were used in the visual laboratory so that we may have the benefit of direct personal knowledge and experience. Lastly, they were selected to include a variety of tests based on different principles in terms of both design and material. Four types are discussed. Those based on the pseudoisochromatic principle (P.I.C.) are in the form of books in which certain designs have to be identified and the ones discussed here are the Ishihara and Dvorine plates, the Tokyo Medical College colour vision tests (T.M.C.), and the American Optical Company, Hardy, Rand and Rittler Plates (AO H-R-R). Secondly, a now well-established test of colour discrimination and colour confusion devised by Commander Farnsworth after the last war, the so-called Farnsworth-Munsell 100-Hue (100-Hue), will be discussed. Thirdly, two anomaloscopes will be described, the well-known Nagel type and the Pickford-Nicolson anomaloscope (P-N). Lastly, some specialized tests designed to detect specific colour discrimination abilities are included. These are the Burnham-Clarke-Munsell Colour Memory Test (B.C.M.T.) and the Inter-Society Color Council, Color Aptitude Test (I.S.C.C.-CAT).

The first part of the paper deals with general aspects of the theory of testing, such as problems of defining and classifying colour defects, followed

by an introduction to the science of colour measurement, whereas the second part will deal with specific tests.

Nature of colour vision

The concept of colour vision is rather complex and underlies a great variety of experiences, which are partly determined by the nature of the physical stimulus and partly depend on the structural and functional peculiarities of the observer's visual mechanism.

If we are to understand 'colour vision' we must be constantly aware that it depends on the nature of the stimulus, the state of adaptation of the sensory organ, and also on the type of observer. As a result a multitude of concepts physical, psychophysical, neurophysiological, and psychological must necessarily be understood and utilized. Perhaps the best way to regard colour vision is as an emergent and subjective experience, dependent at intermediate stages on variables which can be studied and usually verified by objective means. Such an approach will remove some of the mystique associated with such studies and at the same time will allow us to quantify these experiences. Instead of being able to classify colour vision only as defective or normal, a richer and more subtle classification becomes possible, indicating that there is a continuum in ability from those completely devoid of it (achromatism) to those whose colour discrimination is excellent. This conception is infinitely more useful than a dichotomous scheme in deciding which type of discrimination is necessary for a given colour task.

The most fundamental variable in vision is the level of the surrounding illumination which determines the level of visual adaptation and thus the appropriate visual process. Vision is said to be *photopic* if the eye is adapted to daylight, *mesopic* if it is adapted to twilight, and *scotopic* when adapted to minimal amounts of light. Scotopic vision produces only variations in luminosity, whereas photopic conditions also mediate chromatic function, and under these conditions both brightness and colour attributes such as hue and saturation are possible. As such attributes are interdependent it is important in a testing situation that they can be isolated from each other. For example, if it is possible to discriminate between items by observing brightness differences, then even if hue or saturation differences are present no valid inference about colour discrimination can be made. It is therefore necessary to control such variables; but unfortunately this is not always done even in standardized test material, and so it is important for the tester to be aware of these shortcomings and to approach colour vision testing with a critical mind.

Theories of colour vision

Global accounts of colour vision phenomena are provided by the theories of which there are about 40, falling into two different groups, the first type based on observations of certain phenomena that arise from colour mixing and the second based on our subjective awareness of colour experience. The first is stimulus orientated and tries to explain colour vision from outside, whereas the other is subjectively orientated and tries to explain from within. In time these two approaches became crystallized in the so-called trichromatic and opponent-pairs theories. The first approach arose from the ideas of artists and early lithographers and especially from the ideas of Young (1807). They were formalized by Helmholtz (1892) and given an experimental basis by Maxwell (1890). The second approach sprang mainly from the work of physiologists and psychologists and was first formulated by Hering (1878, 1920) but received its mathematical rationalization much later from Hurvich and Jameson (1955). In the 19th century, both approaches were naïve in their explanations of the underlying neural mechanisms, and at present there is still no single theory which gives a satisfactory explanation of all colour phenomena.

The essence of the trichromatic theory is the concept of three *simple* sets of sensory mechanisms (peripheral) whose quantitative characteristics provide a basis for the different type of colour vision discrimination. These are simple response mechanisms involving red, green, and blue and are described as the *fundamentals* which characterize the sensory mechanism as such. Of course it was known even from Maxwell's work that colour matchings with these three primaries cannot reproduce all colours, and thus the system is only an approximation.

The opponent-pairs theory is based on Hering's ideas and states that the visual mechanism involves three *complex* and antagonistic processes, each of which functions as a pair of opposing sensations – one black-white, a second red-green, and the third yellow-blue, the equal stimulation of the two members of a pair always giving rise to a neutral sensation (achromatic or grey). And according to Hering the effect of light depends not only on its physical characteristics but also on the condition of the visual mechanism. For example, the actual hue shifts of individual hues depend on the wavelength of the stimulus, but they also vary to some extent as a function of brightness in the so-called Bezold-Brücke effect, where increased brightness causes shifts in the subjective hues of the surrounding wavelengths towards the four primaries of Hering. Simultaneous contrast again could be explained in terms of antagonistic processes triggered off in areas adjacent to the stimulated zone. This theory was not formally verified by experimentation on the same

scale as were the elements of the trichromatic theory for to Hering's followers it was difficult to experiment objectively on phenomena that were essentially regarded as subjective, though a great deal of experimenting was done on contrast phenomena, and the pseudoisochromatic tests that followed at the end of the last century are based mainly on Hering's idea of opposing-pairs.

Today's versions of these theories take up a compromise position (see Galifret, 1960) for neither alone can satisfactorily explain the complex nature of colour vision. Their exponents are Stiles and Hurvich, who both admit that neither theory is formulated to exclude all the elements of the other, and thus the 80 years or so of feud are slowly being forgotten and differences are being approached rationally.

Modern versions came into being after the last war and were more sophisticated. Their explanations were subjected to more rigorous mathematical treatment; they were based on more experimental work and were expected to bear some relation to the basic neurophysiological elements postulated by the physiologists.

Today as a result of this careful re-examination, for example in the case of the trichromatic theory, a limited number of differently sensitive elements in the retina are assumed to exist (the so-called three receptor type of theory) in which the sensory mechanism is associated with cones and their pigments. This has been strongly supported by a number of studies on cone pigments, where bleaching techniques were used on normal and colour defective observers, by Rushton (1958, 1962) and Ripps and Weale (1963) while Brown and Wald (1964) and Marks, Dobbelle, and MacNichol (1964) experimented on single cones in monkeys and on the human eye using a microspectrophotometer. These studies seem to indicate the existence of three types of cones with different spectral sensitivities although they do not entirely conform with the theoretical maximum sensitivities of the three initial response characteristics postulated by Young and Helmholtz.

Hering's formulation, on the other hand, has received recent confirmation from the work of Svaetichin and MacNichol (1958), De Valois (1960), and De Valois, Jacobs, and Jones (1962), who, experimenting with microelectrodes on goldfish and monkeys respectively, recorded the responses at different levels of the visual system. They have adduced electrophysiological evidence for the existence of red-green and yellow-blue response mechanisms at least at the levels of the ganglion cells and the lateral geniculate bodies.

Then there is the pioneering work of Granit (1947, 1955), some aspects of which could be interpreted in the context of trichromatic and opponent-pairs theories, especially the constructs of

dominators and modulators (broadband and narrow-band threshold responses as a function of wavelength). In the case of photopic dominators these can be thought of as corresponding to Hering's achromatic black-white processes, while modulators may provide cues for wavelength discrimination since their maxima are near the wavelengths that correspond to the primaries of the trichromatic colour vision theory.

Currently available evidence would seem to point to some sort of *stage or zone theory* of colour vision of the type first propounded by von Kries in 1905 and expressed in the contemporary work of Walraven (1966), though earlier work by Müller (1930) and Judd (1958) are all good examples of complex stage theories. This type of formulation suggests that a Young-Helmholtz theory must be valid at the receptor level but that the signals from the receptors are processed in such a way that Hering's opponent pairs colour theory applies at the later stages. For an up-to-date description of these different approaches see Galifret (1960), Hurvich and Jameson (1960), Graham (1965), Judd (1966), and Le Grand (1968).

Classification of colour defects

The knowledge we acquire from such colour vision theories is not directly applicable to the testing situation, for many of the techniques developed over the last 100 years were not formulated from *a priori* principles but in many instances were the outcome of a trial and error approach where a more pragmatic philosophy prevailed. If a given test divided people into those with normal colour vision and those who were defective, then the test was a good one. However, an examination of how colour defects are classified might increase our understanding of how tests were devised and why some are more efficient than others and, again, might help to explain certain difficulties which occasionally arise in the testing situation.

In practical situations it is never satisfactory to classify colour defects on the basis of one theory alone since this cannot account for all the diversity that we find. The consultant very soon finds that to label accurately all the different types that he diagnoses he has to use the language of apparently opposing systems. He must try to combine these systems, and whatever synthesis he finally achieves will depend on many factors. To help us towards such a synthesis different approaches to classifying defects will be discussed, bearing in mind the contributions of the two principal theories.

To show that this is a real problem let us try to explain dichromatism in terms of one theory only. From the standpoint of the trichromatic theory alone, we would say that this defect consists in a loss or absence of one of the fundamentals and

therefore we might talk about 'red blind', 'green blind', or 'blue blind' people. However, such terminology is meaningless for, even under laboratory conditions, these subjects do not talk about absence of sensation in the relevant parts of the visible spectrum. In the case of the so-called 'red blind', the visible spectrum is not seen as something that normal observers see, but is seen without the red component. Such subjects divide the spectrum into two parts, where one part from about 500 nm to the longer wavelength is referred to as yellow progressively fading into a 'dirty' dark colour, while the other half of the spectrum is referred to as blue. In a complex visual situation we find that Hering's idea of opposing-pairs of colours is nearer the observable facts of such a subject's discrimination, for he actually confuses the colours of the pair. He confuses red with green, or orange with yellow or yellow with green, etc. Thus, to say that such a dichromat cannot see red conveys nothing useful if we wish to describe the sort of difficulties he encounters, but if we say that he will confuse all colours between red and green this fits the observed facts, though it still might be true that his particular defect lies in the absence of a red receptor mechanism.

It is with such a problem in mind that this particular section on classification is eclectic in nature. Terminology is used from many theories and occasionally, when classifying predominantly according to one theory, the terminology used belongs to another type. For example, when classifying dichromats (which is essentially a trichromatic approach) we introduce the actual labelling of the various dichromats from language that does not fit the scheme.

The simplest way to speak about colour defects is to refer to them as either congenital or acquired. With the congenital type no morbidity is implied, only deviations, whereas with the acquired type one always looks for a pathological basis. They can be inherited, but the assumption is that at some stage colour vision was normal. At one time it was thought that colour vision tests should be able to differentiate between the two categories.

We may also classify defects in terms of the two major colour vision theories. If Hering's terminology is accepted, we group defects according to the most characteristic pairs of colours which such subjects confuse with white stimulus, thus we have red-green and yellow-blue defects or confusions. If, on the other hand, Maxwell's findings on additive colour mixtures are followed, we can group the defects according to the number of primaries required to reproduce *all* the colours in the colour space, observing that in this theory the irreducible number of primaries is three for the majority of people (normal trichromats). This classification is the most popular in English-speaking countries and also

provides us with the best known and most frequently used terminology.

If under photopic conditions only one primary is needed to match any spectral colour we talk about *monochromatism* or achromatopsia. Such subjects have no colour discrimination but only the photopic luminosity function. Occasionally this results in a scotopic rather than a photopic relative luminosity function. To distinguish between these two types of defect we define them as either cone or rod monochromats; their frequency is minimal.

If, on the other hand, not more than two primaries are required to match the whole colour gamut we talk about dichromatism. These are the defects most often referred to when we talk about colour blindness. Four types are distinguished and they may be grouped in pairs, two along the red-green confusion axis, that is, *protanopia* and *deutanopia* (literally meaning 'first' and 'second' defect and named so by von Kries so that no inferences about the physiological nature of the defect would be made from nomenclature), and two along the yellow-blue axis, *tritanopia* and *tetartanopia*. The frequency of the two red-green types is about 1% for each among men and is even lower among women. Data for the incidence of *tritan* defects is less certain and least of all for the fourth type, tetartanopia, which some think is not congenital in origin at all although it is thought to be so by Müller (1924) and by Judd (1949). This will be discussed again in a later section dealing with colour measurement.

Lastly, if three primaries are necessary, but the ratios of any of these are appreciably different from those of normal observers, we have the so-called *anomalous trichromats*, who are referred to as deuteranomalous, protanomalous (terms coined by Nagel (1907)) or tritanomalous. The incidence of these anomalies is higher than the incidence of previous defects, being 4% to 5% for the deutan type and 1½% for the protan. No reliable data exist for the third type. Their visual characteristics will be discussed when the various tests are described. A more detailed account of the red-green types of defect will be given later, especially in the section on anomaloscopes, where what is meant by anomalous defects can most easily be understood.

Classification of colour defects can also be according to prediction of which is the impaired part of the visual mechanism. Such a classification was propounded in 1924 by von Kries, who recognized three systems, the so-called absorption, alteration, and reduction systems. The essential characteristic of the *absorption system* is that the retinal mechanism is normal and colour deviation is due to pre-receptor changes which are caused by changes in either the cornea or lens. It is a *non-selective* absorption system if the modification of the retinal illumination is identical for all monochromatic

radiations of the visible spectrum, and *selective* if this modification varies with wavelength. In the *alteration* and *reduction* systems the anomaly is attributed to a defect in the retinal mechanism. The alteration system is characterized by the non-acceptance of certain of the normal colour matches, best seen in the anomaly quotient of anomalous trichromats (see section on anomaloscopes). In the reduction system we find the acceptance not only of the colour quotients of normal subjects but also of the anomalous. Usually the reduction system is defined as a form of dichromatism. In this system there are only two fundamental response curves instead of the normal three curves, and this is also sometimes called the collapse system or the collapse of the normal fundamentals (see section on colour measurement). For further reading on this topic, consult Pickford (1951, 1958), Verriest (1964), and Kalmus (1965).

Changes in colour vision

Congenital colour defects, frequently quoted as being the only form of faulty colour vision, have been admirably discussed in a recent book by Kalmus (1965) with particular reference to the way such defects are inherited, and from an excellent and readable paper by Wald (1966) we may gain considerable insight into the molecular and genetic mechanism underlying this type of vision. An extensive account of the frequencies of these defects throughout the world is given in a book by Kherumian and Pickford (1959). But changes leading to losses in discrimination, which might be as extensive as or similar to the congenital types, can be brought about by processes such as normal ageing or by pathological conditions of the eye itself or as a by-product of some systemic diseases. Though our knowledge of the extent and frequency of such losses is recent, it is important to stress their existence since ageing processes affect performance on almost all colour vision tests, and losses due to disease are much more frequent than was previously suspected or admitted.

Colour vision changes due to age

If studied within an age continuum, that is, from childhood to old age, it can be seen that there is a growth of function up to the age of 20 and a decline after about 30 years. Such age variations were observed in the late thirties by Smith (1943) and by Tiffin and Kuhn (1942) but their significance has been accepted only within the last decade, and recent work shows that such variations can be found to apply to all types of discrimination. Gilbert (1957) showed this to be true for performance on the C.A.T., Lakowski (1958, 1962) measured it by colour vision testing on the anomaloscope and a number of

P.I.C. tests, Verriest, Vandevyvere, and Vanderdonck (1962) did so for the Farnsworth-Munsell 100-Hue scores, and Ruddock (1965) for discrimination on Wright's colorimeter.

Because of its importance, the age variable will be included in the discussion of the various tests, and the extent of losses will be properly assessed, so that the information will be available to help us to distinguish congenital defects from the effects of senescence.

How the age effect comes about is not properly understood and its causes have yet to be fully established. Some of the minor changes in colour vision can quite clearly be attributed to the ageing of the lens and/or macula (Said and Weale, 1959; Lakowski, 1962; Weale, 1963). The more extensive losses might be attributable to the fact that after 30 there must be increasing numbers of people who, as well as showing the normal changes, have minor pathological conditions, which probably accelerate the senile degenerative processes so that by, say, 50 the use of the term 'normal senile changes' becomes of questionable value. Recent studies on specific acquired dyschromatopsias which have shown that in many cases pathological conditions give rise to colour vision losses similar to those of the very old, a kind of early ageing effect, tend to confirm this. The possibility of pre-clinical conditions affecting colour discrimination, and in all probability aggravating the 'ageing effect', must not be ruled out.

Acquired dyschromatopsias or acquired colour losses
This is the name given to those losses which are not congenital where it is always assumed that at some stage the subject's colour vision was normal. They may be hereditary or traumatic in origin or due simply to pathological senile degenerations of vision. It is perhaps useful to point out some of the differences between congenital and acquired dyschromatopsias. For example, we know a great deal about the heredity of congenital defects but very little about acquired ones. Although the visual mechanism underlying congenital defects is not fully understood, with acquired dyschromatopsias we can infer where the impairment lies from a knowledge of pathology. It is axiomatic that in congenital defects no morbid conditions exist but pathological conditions are always assumed to exist with acquired ones. In congenital defects the results of colour vision testing are stable and predictable, the axis of confusion is definite, colour naming always shows classical confusions (*e.g.*, red or purple and brown mistaken for grey or green by the red-green defectives), and the defect is bilateral. In acquired dyschromatopsias, on the other hand, results are unstable, progressively deteriorating or occasionally improving if the underlying pathological conditions are reversible. The axes are not so definite, there is

usually good colour naming, and there are variations between eyes or even between parts of the retina (on axes of confusion see section dealing with colour specification).

The earliest studies in this field were made in the second half of the nineteenth century during the so-called German period, when it was shown that acquired dyschromatopsias, though not so clear cut, could be subdivided in the same way as congenital dyschromatopsias into abnormal trichromatism, dichromatism, and monochromatism. It was Koellner (1912) who postulated that there are dyschromatopsias of the red-green and yellow-blue axes as well as some without any apparent axis. It was also established that lesions in the outer layer of the retina lead to yellow-blue defects, whereas lesions in the inner layers or in the optic nerve lead to red-green defects. For further information about causes, manifestation, and frequency of acquired dyschromatopsias see Zanen (1953), Francois and Verriest (1957, 1961), Verriest (1963, 1964), Jaeger (1956), Grützner (1962), and Cox (1961 a and b). It would be interesting to speculate on their incidence in relation to the general population. At present no such data exist but it would not be exaggerating to say that about 5% are as severely affected as the existing 8% to 10% of congenital colour defectives.

Ageing and pathology as causes of colour vision losses have been deliberately introduced because they are sometimes ignored or their effects minimized. But in practical situations, where the question of upgrading older workers often arises, such losses might be regarded as congenital and, if only recognition of primary colours is involved, a worker might be unnecessarily penalized. On the other hand, these losses might go unnoticed, especially where P.I.C. tests alone are used for detecting congenital red-green defects. Nevertheless the discrimination losses concomitant with these changes are usually large enough to have an effect, especially where good colour discrimination is required or where yellow-blue discrimination is very important, as for example in the printing trades or in textile colour matching.

Basic elements in colour vision tests

To understand the merits and attributes of *specific* tests we must consider the basic elements in colour vision tests as a whole. The following classification is an attempt to show what elements underlie many apparently different tests. It is perhaps advisable to read this section twice, once before a more detailed account of the tests is given, and again after reading the second part of the paper. The aim is to indicate similarities and differences in terms of colour vision and stimulus characteristics in tests that are apparently of the same type.

Mode of appearance

In an ideally constructed test 'colour' should be confined to the non-spatial, non-temporal attributes of colour perception, to those aspects of discrimination which involve least interpretation in terms of past experience and are unaffected by intelligence or personality. Yet when colour appears in tests even in its simplest form in the so-called 'modes of appearance', it is always at least in a spatial or temporal context. These are the most fundamental elements of a test and three such modes are distinguished – surface, volume, and film. When colour is perceived as belonging to a surface, it is said to be perceived in the *surface mode*, and the majority of colour vision tests such as the Ishihara, Dvorine, 100-Hue, and C.A.T. fall into this category. When colour is perceived through the bulk of a uniformly transparent substance it is said to be perceived in the *volume mode*. This is not important in the realm of colour vision testing. When *film mode* is referred to, the implication is that the stimulus object is not distinguishable as an object, for example where colour seems to fill a space behind an aperture. This is what happens in most colorimeters and also in anomaloscopes. The film mode of appearance is considered the simplest of all modes, in that it is the least removed from pure sensation, whereas with surface mode the complex attributes of full visual perception such as form, size, glossiness, texture, glitter, etc. can never be estimated. For further information see Judd (1958), Burnham, Hanes, and Bartleson (1963).

Complexity of visual task

Some tests are simple and some complex in relation to the cognitive and perceptual involvement required. The degree of cognitive involvement needed by most P.I.C. tests is surprising, for subjects must know numerals (usually arabic) or recognize shapes (usually geometric). Then in colour discrimination tests, such as the 100-Hue, the concept of a 'colour series', say from red through orange and yellow to green, plays an important role in the ability to perform the test well, and this depends not only on the physiology of vision but also on the maturation of cognitive processes.

Some tests, on the other hand, present relatively simple cognitive and perceptual tasks. They require an average amount of intelligence and depend mostly on the neurophysiological state of the subject's visual system. Anomaloscopes fall into this category. Also in such instruments the colour match achieved is relatively independent of the state of adaptation. Though the colour naming will change, the ratios of the primaries in such a match will remain the same over a large range of adaptation levels, provided that the general luminosity value of the matching field is above the scotopic level. Thus

such an instrument can be used at mesopic as well as photopic levels of illumination without essentially altering the relative scores. This does not apply to tests based on the P.I.C. principle, where adaptation to the level of illumination and the colour temperature of the illuminating source are important factors controlling the type of response and so the scoring on the tests.

Complexity of stimulus

In physical terms, the stimulus in P.I.C. tests is complex. The resultant colour depends not only on the energy distribution of the illuminant but also on the peculiarities of the pigments involved, which can have 'broad' or 'narrow' band characteristics. Each plate in such tests is composed of many differently coloured dots (four to 10 or more) each with different reflectances, which make up the figure and background. Anomaloscopes also vary in complexity. In the Nagel, monochromatic stimuli are used, whereas in anomaloscopes that use filters, rather broader transmissions are used with consequent reduction in the purity (saturation) of the stimulus.

Colour vision tests differ in another very important aspect, that is in the degree of metamerism or isomerism that is present in the matching situation. These terms were adopted by Ostwald (1931-33) to describe the phenomenon exhibited when two or more surface colours seem to appear alike under one illuminant such as daylight but do not match under another illuminant such as an incandescent lamp. Although these terms were coined only about 40 years ago the phenomenon was by no means unknown before. Studies of extreme metameric pairs, in which mixtures from two different parts of the spectrum are set up to match another spectral mixture in colour, have yielded the most valuable information about the properties of the average normal eye (the so-called Grassman Law). In this sense metameric pairs, as they are called, have to be seen as stimuli with the same subjective colour but different spectral composition.

Thus when colours look alike, whether they belong to a surface or come from a source, they do so because *either* they are physically or spectrophotometrically identical (this is called an 'isomeric match') *or* they are physically dissimilar, *but*, because of the peculiar nature of colour vision, they match, that is, they give rise to identical sensations (this is called a 'metameric match'). Another way of stating this fact is to say that they match colorimetrically though not spectrophotometrically. If members of a pair are of very different spectral composition, and they are still perceived as identical, they are said to exhibit a high degree of metamerism. In anomaloscopes and colorimeters the colours to be matched are essentially in a metameric relationship, that is the two fields evoke the same sensations

but are spectrally different. In such tests we measure the degree of metamerism that can be accepted by an individual, and this form of colour vision testing is one of the most efficient as it measures directly differences in sensitivity of the individual elements of the visual system, for example, at the retinal level, say the property of 'red' to 'green' 'mechanisms'. In many tests where colour discrimination of surfaces is called for, the differences are essentially isomeric in nature and thus the task consists of perceiving differences in brightness or saturation only. In some P.I.C. tests the surface colours used might also be in a metameric relationship (see Lakowski, 1965 a and b), and there is one test, Glen's Colour Rule, which measures metamerism by means of surface colours, where this is deliberately introduced.

Retinal subtense

Colour vision tests may vary in the size of the image subtended at the retina. In the Nagel anomaloscope a $3^\circ 15'$ subtense is used, while on the P-N anomaloscope it is $1\frac{1}{2}^\circ$, and the diagnosis of the extent of a particular subject's acquired defect may not be the same for the two instruments. Where tests involve surface colours, differences in subtense are even greater. The little coloured dots constituting the figure or background of a P.I.C. plate, viewed at 30 or 40 cm., may subtend from as little as $40'$ at the cornea for the T.M.C. test to as much as $1\frac{1}{2}^\circ$ for some of the dots in the Farnsworth F2 tritan plate. In the 100-Hue each individual cap subtends $2^\circ 10'$, and in the C.A.T. the size of each plastic chip is approximately 35×45 mm., giving a subtense of 3 to $4\frac{1}{2}^\circ$ at normal viewing distance.

Considerations of subtense are important, because colour vision up to 2° is essentially different from colour vision at a larger subtense. When assessing results of a test it is vital to know how much of the retina was stimulated since this is an important factor in a comparison of results of different tests or even between different P.I.C. tests. The aetiology of deterioration due to age or disease might be different for the parafoveal as distinct from the foveal regions, and results of tests designed to detect the same anomaly might be contradictory if the subtense varied to any extent.

Enough has been said for us to realize that a critical evaluation of basic elements of each test is vital especially when more than one test is used, say for difficult cases. The apparent discrepancies between test results might be due to the fact that the basic elements in tests are different rather than to the unreliability of a subject's responses (see Lakowski, 1968 d).

Administration

The secret of good testing procedure lies in being

able to elicit valid and reliable results. Whether this is successful will depend to a large extent on whether the instructions for administering the particular test have been followed, and on whether the relevant information has been extracted. In most situations subjects are quite helpful but, when they know they are defective, when they are suspicious of testing, or when the subjects are either young children or older people, it is necessary to establish proper rapport or the results may be invalid even using the best tests.

Now let us look at the elements unaffected by the human factor. *Valid* results depend not only on the test being 'correctly' constructed (this will be discussed later) but also on whether the specified testing conditions have been followed. The user can do little to improve the test but he can render his results invalid if he alters the administration practice. Most tests give explicit instructions which must be followed. *Reliable* results can be obtained only if the testing situation remains constant from session to session, and testing procedure must be similar for all test users if the results are to be comparable. If conditions differ to any extent, valid and reliable results can still be obtained for each individual testing situation, but no direct comparisons can be made. Discrepancies in administration are most likely to arise over the timing of the test, the viewing geometry, and the type of illumination used.

If timing is critical it is specified. In the 100-Hue, for example, subjects are told that the test is timed, and it is found in practice that, unless such instructions are given, some subjects with good discrimination look upon it not as a test of sensory discrimination but as one of aesthetics and are prepared to take a very long time over it with no appreciable improvement in scores. Timing is also very critical in the Colour Memory, but not in the Colour Aptitude Test. No such instructions are given for most P.I.C. tests, but if comparisons of misreadings are to be made between large-scale studies, most research workers limit presentation time for each plate to 5 seconds, and it is therefore good practice to do this for vocational testing also.

Longer presentation only encourages the use of extra cues, the most important of which are found by changing the angle of viewing. As many P.I.C. tests are printed on glossy paper it is vital to see that the viewing geometry is correct if we are to eliminate the *specular component* which usually enhances any contrast between figure and background. The effect of this can be observed in the very natural habit that colour defectives have of slanting the individual cards, so that the specular component appears. In all testing, therefore, the viewing conditions should be those recommended by the International Commission on Illumination (C.I.E. 1931), where the surface of a test is illuminated at an angle of 45° and

viewing is perpendicular to the surface (short notation for this viewing is $45^\circ/0^\circ$). The illuminant and viewing conditions can be interchanged without any effect (Fig. 1).

Nearly all tests employing surface colour require standard illumination in relation to both the relative energy distribution and intensity. The energy content of the standard light source should resemble that of northern light, but as this varies even during daytime from about $6,000^\circ\text{K}$ to $15,000^\circ\text{K}$ or more (as measured by the colour temperature criterion), an agreed colour temperature such as the C.I.E. standard illuminant 'C' of $6,700^\circ\text{K}$ or the newly proposed C.I.E. illuminant 'D' with colour temperature of $6,500^\circ\text{K}$ should be the rule. In the past it was always difficult to obtain an illuminant 'C' and therefore approximations to such a standard source were acceptable; for example, the Macbeth Lamp 40 BK $6,500^\circ\text{K}$ in the United States of America or modified tungsten lamps run at a colour temperature of $2,850^\circ\text{K}$ (called C.I.E. illuminant 'A') corrected to the temperature of northern light by filters such as Chance Pilkington OB 10 or Corning Daylight filters (Judd, 1961). Recently, however, artificial daylight sources have become easily available, some of which are manufactured in this country by Leslie Hubble Ltd., of Leicester. These have colour temperature of $6,500^\circ\text{K}$ and comply with the B.S. 950 (1966) or the C.I.E. (1966) illuminant 'D' $6,500^\circ$. On no account should ordinary tungsten or fluorescent light of the 'warm tone' type be used in place of a standard illuminant since these are run at a temperature of $3,000^\circ\text{K}$ or less and would alter the intended colour differences (see section on colour measurement). In other words, the intended colour contrast, say, between figure and background in P.I.C. plates would be exaggerated, the recognition of correct numerals would be easier, and even congenital red-green defectives might be able to read them. Again such low colour temperatures would appreciably alter the spacing of the blue caps in the 100-Hue test.

Lastly, it is very important that the intensity of illumination should be correct. If it is too low the test might become more difficult and the result would

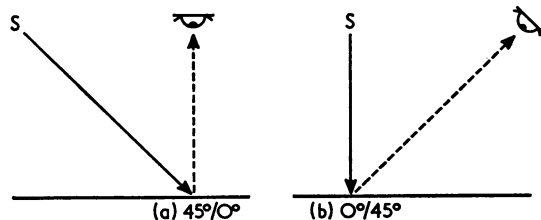


FIG. 1. C.I.E. viewing geometry; the illuminant (S) and viewing conditions are interchangeable.

be poorer scores. If it is too high, then unnecessary expenditure is incurred, as sources run at high intensity levels are costly to buy, to run, and to maintain. Every test has its optimum level of illumination, but the general rule is that the level of adaptation should be above the upper limit of mesopic vision, that is, the luminance level of the viewing field must be above 10 candelas per square meter (cd/m^2) with the natural pupil. As this is the minimum requirement, a more useful level is reached when luminance is twice or three times higher.

As the amount of light arriving at the receptors is not strictly a constant fraction of the intensity of a source, it is essential to calculate the *retinal illumination* in all visual testing. This must be at least 70 to 100 *effective trolands* when testing with colorimeters, and when cognitive tasks are involved it is desirable that it should be between 150 and 300 trolands. The following example illustrates how these values have been arrived at (see also Fig. 2). The photometric units used are metric, but British equivalents are given where relevant.

To reach, say, the 200 effective trolands level of retinal illuminance, the initial illuminance provided by the source should be about 300 lux (lumens per square meter), *i.e.*, 28 foot-candles (lumens per square foot) at the test surface. The resultant *luminance* level of the viewing field will be about $93 \text{ cd}/\text{m}^2$ if a perfect white diffuser is being viewed. Note that this is about one third of the original illuminance because the perfect diffuser now becomes a secondary source with the property of sending back all incident light, but equally in all directions (according to Lambert's Law) so that the amount of luminance of its surface is only $1/\pi$ of the original illuminance. In addition, as all surfaces

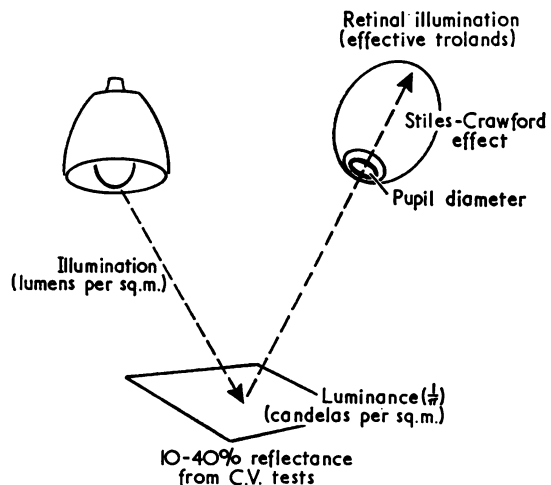


FIG. 2. Stages determining final retinal illumination.

absorb some incidental light, and pigments employed in colour tests absorb between 60% and 90%, only a fraction of the luminance, about one quarter – say between 20 and 30 cd/m²—is effective in producing retinal illuminance.

The actual level will be determined by another factor besides the luminance of the surface viewed, namely the pupil aperture. In any given subject the size of this will depend on age, state of health, excitation, etc., but in our case it is dependent only on the general luminance level of the field which gives rise to a pupil diameter of 3.4 mm. Thus retinal illuminance is the product of pupil area (S) in mm.² and the given luminance (L) in cd/m.² In this case L.S. is 20 or 30 cd/m.² × 11.5 mm.² (the approx. pupil area), or 230 to 245 *trolands* (so one *troland* corresponds to a luminance of 1 cd/m.² viewed through a pupil area of 1 mm.²). If in a given situation pupil size and luminance are known it is easy to calculate the retinal illuminance.

However, because of the Stiles-Crawford phenomenon the actual retinal illuminance will be slightly less than the simple product of L.S. This is due to a rather curious property, the relative directional efficiency of the retina, which is more sensitive to rays striking it perpendicularly than to those which are slightly oblique (Stiles and Crawford, 1933). This means that rays entering the pupil centrally are more efficient than those entering near the pupil margin, so that the effective pupil area must be reduced in this case to 8 mm.² In this case then the effective *retinal illuminance* is between 160 and 240 *effective trolands*. For further discussion on this topic see Le Grand (1968, Chpt. 5 with particular reference to the table of pupil diameter and areas at various luminances), Grossman (1967, Chpt. 3), Sidowski (1966, Chpt. 7), Graham (1965, Chpt. 1), and Crawford (1968).

Colour specification

Normally when talking about the coloured world our descriptions are subjective, which leads to a richness in linguistic expression seen in the names found on designer's cards such as 'subdued persian golden brown', etc. Though doubtless meaningful to the individual who coined them, such colour names are not readily identifiable by others. Because it is difficult to know what a specified colour really looks like, about 30 years ago a system of description in terms of the physical attributes of the stimulus was accepted, which has transformed our methods of describing colour. This can no longer be subjective or man-orientated only but must take account of the precise nature of the stimulus and must accept a *Standard Observer* as a reference point in colour perception. The system allows us to measure colour independent of the colour vision deviations

found among normal observers. When specifying colour we must include not only the spectral composition of the radiant energy reaching the observer's eye but also the visual properties of the observer himself.

Such specification is becoming common practice in the industries that deal with colour, and there are several reasons for including a description of them. First, a great deal of our understanding of test design depends on being able to give exact specifications of the colours used. Secondly, it is best to evaluate the degree of difficulty involved in a test from the colour differences in it, and these in turn can be calculated from the objective data supplied by colour measurement. Lastly, certain aspects of defective colour vision such as colour confusion can best be illustrated by reference to this colour space.

In the colour theory section it was pointed out that a normal observer requires at most three primaries to match a given colour: red, green, and blue if matching is done with lights; red, yellow, and blue for paints; and magenta, yellow, and cyan in lithography or photography. Three primaries are the irreducible minimum, and it has been concluded from this and other studies that normal colour vision is trivariant, that is, that not more than three variables are required to determine it. It is therefore possible in a *mathematical system* to express colour by three numbers. The modern science that deals with this type of specification is called *colorimetry* and the particular system of specifying colours described here is the one adopted by the International Commission on Illumination (or the C.I.E.) in 1931. In this system three imaginary primaries X Y Z, embodying the characteristics of the real R G B (red, green, blue) primaries, completely specify the various mathematical functions in the system, and specification of a given colour involves finding the relative amounts of the three primaries that are required to match that colour under a standard illuminant. It would be a mistake to infer from the use of the number 'three' that this tends to confirm the Young-Helmholtz theory of colour vision. The C.I.E. system shows only that a three-dimensional method of specification is both workable and useful. At most it gives one way of interpreting the fact of trivariance (Le Grand, 1968).

Figure 3, illustrating the geometrical relationship between the XYZ and RGB systems, shows the mathematical primaries as the three apices of a right-angled triangle within which the 'real' colour space is inscribed in the form of the so-called *spectrum locus*. The numbers on this line denote the positions of certain wavelengths in nanometres (nm) and the positions of the three real colour primaries (RGB) are also shown, as are the loci for the C.I.E. illuminants 'C' and 'A'.

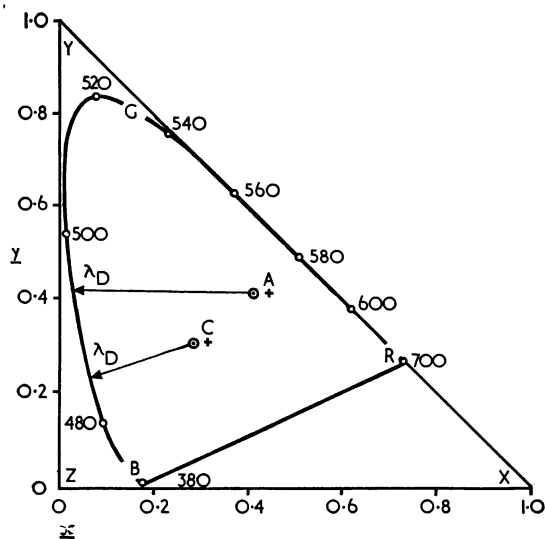


FIG. 3. Chromaticity diagram for the XYZ system (after Le Grand, 1968). RGB denotes positions of the real primaries, while letters A and C mark positions of C.I.E. standard illuminants. The co-ordinates x and y are marked at intervals of 0.2 unit.

Colour measurement

The mathematics involved in transforming measurements obtained from colour mixing with 'real' primaries into the XYZ system is complex (see Le Grand (1968) or Graham (1965)) but at present all that is needed is a knowledge of how colour is specified in the C.I.E. system, for here the tristimulus values (XYZ) for any colour are simply the product of the physical or spectrophotometric quantities of the stimulus and the colour matching quantities required for it by the 'average eye'.

To make this system meaningful without involving too much theory, an example of colour measurement (Fig. 4) is worked out for the blue cover of this journal (specifically for the July 1968 issue). This is in the form of a flow diagram showing how we arrive at the tristimulus values for the cover. The following explanation should be read in conjunction with the six diagrams in this illustration.

The tristimulus quantities XYZ for this blue cover as matched by the 'average eye' under illuminant 'C' are indicated in Fig. 4 a-f and they were derived as follows:

First, the initial values of the distribution coefficients of the Standard Observer (as shown in Fig. 4a) were modified by the energy distribution of illuminant 'C' ($E_{c\lambda}$) shown in Figure 4b. These coefficients simply represent the amounts of the C.I.E. primaries required by the Standard Observer (mean of 17 measurements by actual observers

obtained by Guild (1925-6) and Wright (1928-9) and accepted by the C.I.E. in 1931 as the visual standard), to match narrow band wavelengths (of equal energy) throughout the visible spectrum. Then these so modified distribution coefficients are multiplied by the spectral reflectance factor of the surface (ρ_λ) (Fig. 4c). If the areas under the three curves in Fig. 4 d-f are summated, we obtain numerical values for these. Table 1 shows such values obtained for the April and July 1968 issues of the Journal under two C.I.E. illuminants 'C' and 'A' and indicates the amounts of XYZ required to

TABLE 1
AMOUNTS OF TRISTIMULUS VALUES XYZ
REQUIRED TO MATCH BLUE COVERS OF APRIL AND
JULY 1968 ISSUES OF *Brit. J. industr. Med.* UNDER
ILLUMINANTS 'C' AND 'A'

Issue	Tristimulus Values					
	Illuminant 'C'			Illuminant 'A'		
	X	Y	Z	X	Y	Z
April ..	304.0	321.5	427.0	195.0	190.2	83.9
July ..	321.3	337.5	446.0	203.7	197.8	87.3

match these covers under standard northern and tungsten filament lights. Note that under the more bluish illuminant relatively more units of Z are required for matching both covers than for matching them under artificial light. Note also that under both illuminants there are slight differences in tristimulus values for the April and July issues which will be commented on later.

Using these results we can also find the *relative brightness* of the blue covers from the Y tristimulus value by comparison with the standard white. Let us look again at Fig. 4a, paying particular attention to the shape of the \bar{y} distribution coefficient. In the C.I.E. system this function has been deliberately chosen to be identical to the characteristics of the relative luminosity efficiency curve of the human eye (V_λ photopic - see section on Nature of Colour Vision), and thus the value of Y when expressed as a percentage of the Y of the comparison standard used in reflectance measurements magnesium oxide (MgO) gives the brightness or luminosity factor for

this surface $\left(Y\% = \frac{Y \text{ sample}}{Y \text{ MgO}} \right)$. In our example, this is as shown in Table 2.

Thus under both illuminants the covers will have almost equal luminosity factors, looking slightly brighter under illuminant 'C' than under illuminant 'A'.

In the C.I.E. system, therefore, the tristimulus

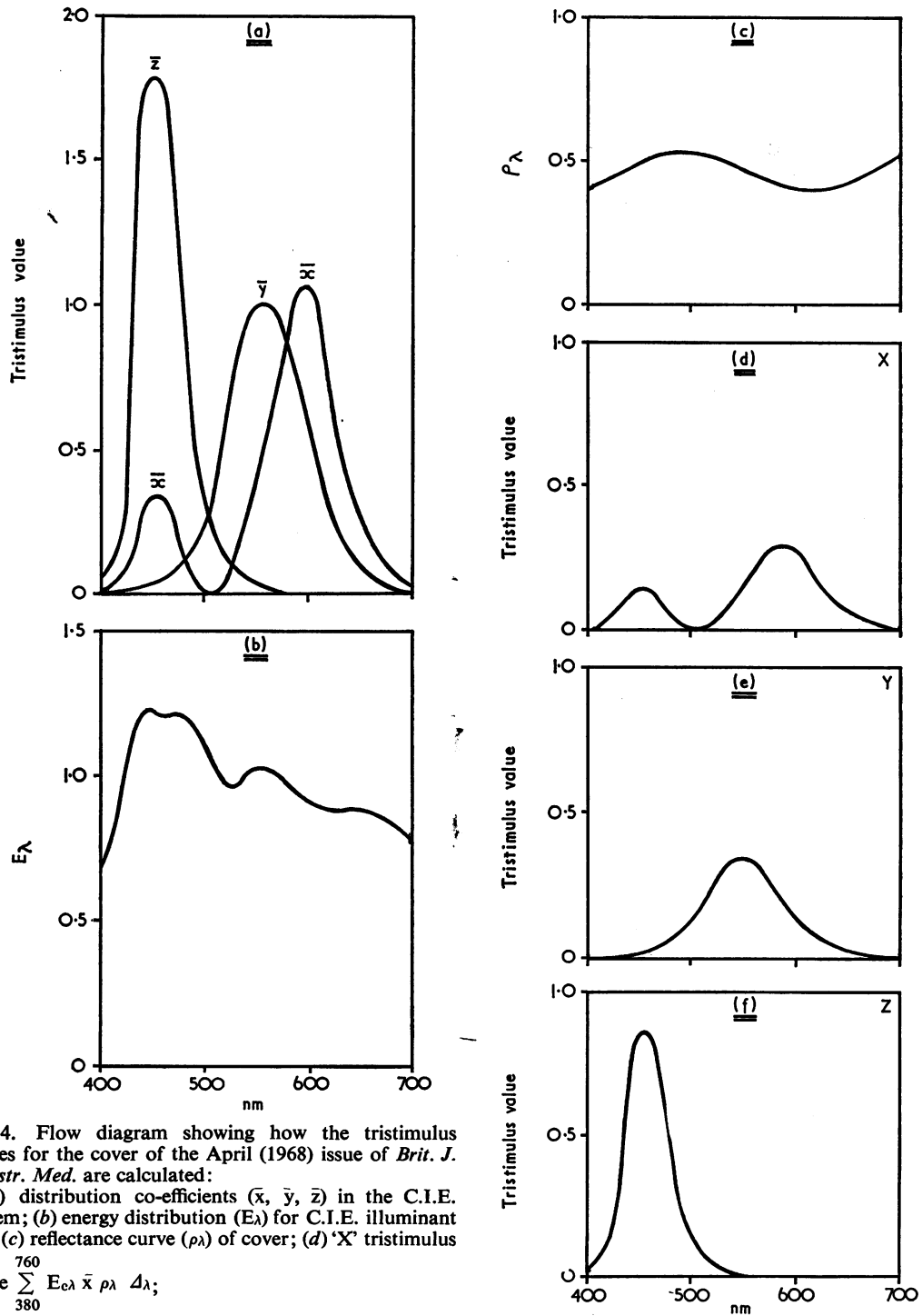


FIG. 4. Flow diagram showing how the tristimulus values for the cover of the April (1968) issue of *Brit. J. industr. Med.* are calculated:

(a) distribution co-efficients (\bar{x} , \bar{y} , \bar{z}) in the C.I.E. system; (b) energy distribution (E_λ) for C.I.E. illuminant 'C'; (c) reflectance curve (ρ_λ) of cover; (d) 'X' tristimulus

value $\sum_{380}^{760} E_{c\lambda} \bar{x} \rho_\lambda \Delta\lambda$;

(e) 'Y' tristimulus value $\sum_{380}^{760} E_{c\lambda} \bar{y} \rho_\lambda \Delta\lambda$;

(f) 'Z' tristimulus value $\sum_{380}^{760} E_{c\lambda} \bar{z} \rho_\lambda \Delta\lambda$.

TABLE 2
LUMINOSITY FACTORS OR Y% FOR THE
COVERS OF APRIL AND JULY ISSUES OF THIS
JOURNAL

Issue	Tristimulus Value 'Y'					
	Illuminant 'C'			Illuminant 'A'		
	MgO	Blue cover	Y%	MgO	Blue cover	Y%
April ..	713.0	321.5	45.09	434.0	190.2	43.82
July ..	712.0	337.5	47.43	434.0	197.8	45.56

values XYZ and the Y% are sufficient to describe the stimulus colorimetrically.

The C.I.E. chromaticity diagram

In practice, however, the absolute values of XYZ are not usually given (as these vary from instrument to instrument). What really matters is the relative quantities of each that are necessary for a given colour match. It is more usual to express each tristimulus value as a proportion of the sum of the three values, and these ratios are then termed the C.I.E. co-ordinates, symbolized in the letters *x y z*. They are derived as follows:

$$\text{Co-ordinate } x = \frac{X}{X + Y + Z}$$

$$y = \frac{Y}{X + Y + Z}$$

$$z = \frac{Z}{X + Y + Z}$$

The reader may confirm the co-ordinates of the blue covers from the tristimulus values given in

TABLE 3
COLOUR SPECIFICATIONS FOR THE 'BLUE'
COVERS OF THE APRIL AND JULY 1968 ISSUES
OF *Brit. J. industr. Med.*, IN TERMS OF C.I.E.
CO-ORDINATES *x* AND *y* AND THE Y%;
SPECTROMAT FS2. (0/45° GEOMETRY,
MgO STANDARD)

Issue	C.I.E. co-ordinates					
	Illuminant 'C'			Illuminant 'A'		
	<i>x</i>	<i>y</i>	Y%	<i>x</i>	<i>y</i>	Y%
April ..	0.289	0.305	45.1	0.416	0.405	43.8
July ..	0.291	0.306	47.4	0.417	0.405	45.6

Table 1. These co-ordinates are shown in Table 3.

The co-ordinates indicate very simply the chromaticity of a given surface colour in the C.I.E. chromaticity diagram (or the mixture diagram). Note that in Table 3 only values for two chromaticity co-ordinates are quoted for each colour. Since, by definition, the sum of the chromaticity co-ordinates is unity ($x + y + z = 1$), the value of the third is simply arrived at. Thus colours are most frequently specified by giving any two of the co-ordinates (*x* and *y*) plus the luminosity value Y%. In all subsequent diagrams showing positions of colours used in the various tests the actual co-ordinate values are not specified but approximate readings correct to the second decimal place can be read off from the scale, for co-ordinate *x* (*i.e.*, the ratio of the tristimulus value X to the sum of all the three tristimulus values X + Y + Z) on the abscissa, and for co-ordinate *y* (ratio of Y to this sum) on the ordinate.

It should be remembered that each complete diagram, such as is shown in Figs 3, 5, and 6, represents only one plane of the colour solid taken at right angles to the vertical white and black axis, where the spectrum locus marks the boundary of all real colours in that plane. Sometimes only a part of this diagram is shown but in such cases the co-ordinates *x* and *y* allow us to place this section in relation to the total C.I.E. colour space. In addition, the position of the white illuminant and any part of the spectrum locus can be used as reference points. The C.I.E. diagram does not indicate brightness values (Y%).

Other properties of the C.I.E. diagram The diagram can be looked on as a map of colour space, and because of this it can be used as a convenient shorthand for explaining certain aspects of colour relationships. The most obvious comparison is between the colours used in one particular colour vision test and those used in other tests. Once the position of a colour is specified in the diagram, inferences about its hue or saturation can easily be made.

In the C.I.E. system, the concept of *dominant wavelength* (λ_D) corresponds most closely to the psychological attribute of colour called *hue*. The λ_D of any colour is determined by drawing a straight line between the position of the reference standard white and the position of that colour and extending this line until it intersects the spectrum locus. This point of intersection will indicate which part of the spectrum will match the given colour when mixed with a neutral standard (*e.g.*, white). In Fig. 3, for example, notice the loci of the blue covers (marked as ●) in relation to illuminant 'C' and 'A'. The lines drawn through them to the spectrum locus show that λ_D for the April cover is 483.5 nm when seen under

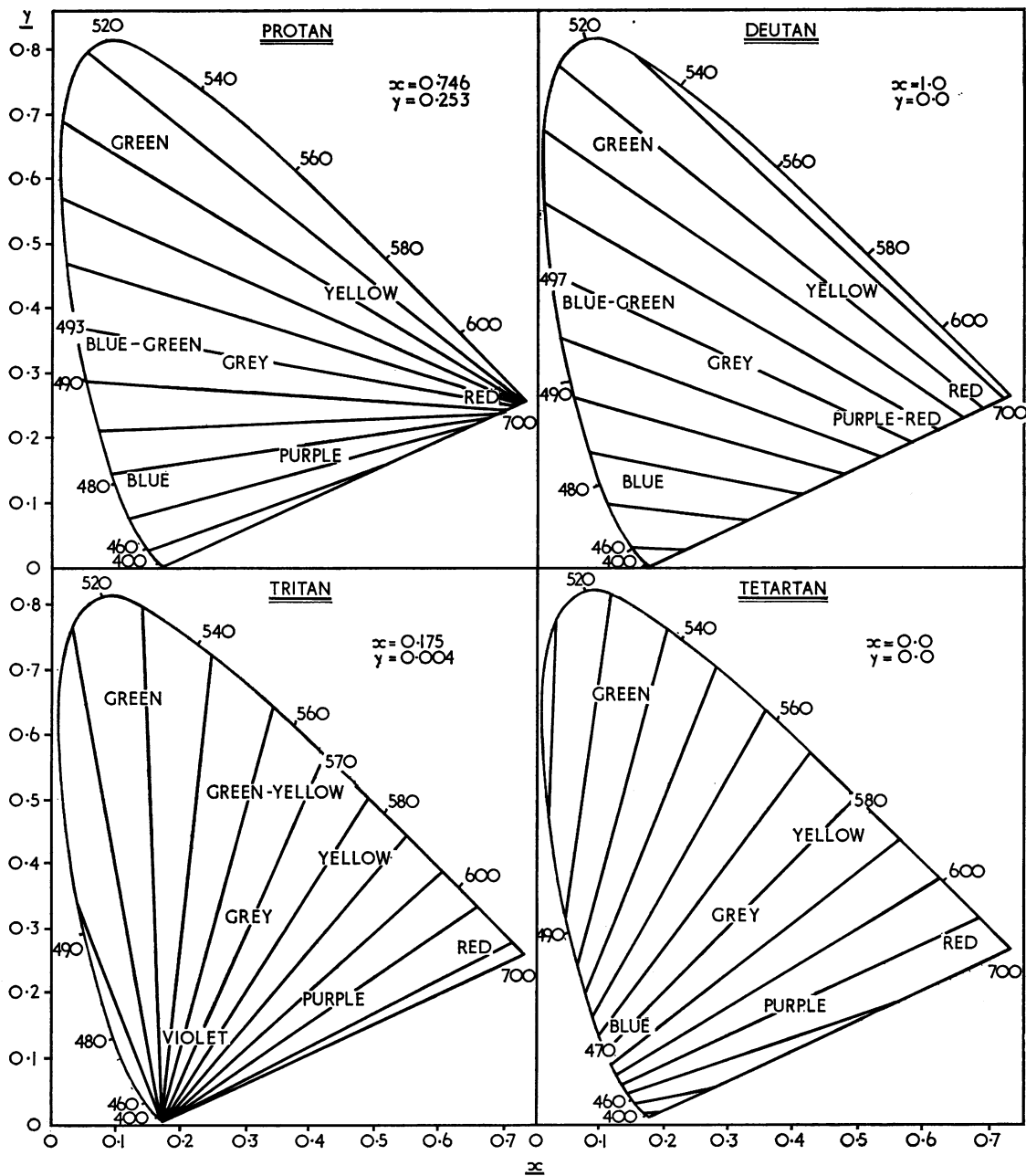


FIG. 6. Confusion loci, centre of confusion, and neutral axes for dichromats.

and they may be looked on as theoretically representing the positions of the missing fundamentals. For each type of dichromat the confusion locus which passes through the neutral part of the colour space (grey or the standard illuminant) and intersects the spectrum locus is called his *neutral axis*. This shows very simply what colours in the C.I.E. space a given dichromat will be unable to distinguish from neutral and which part of the spectrum he will match with the standard white. Dichromats have no difficulties in finding such 'neutral points' in the spectrum, whereas the anomalous and normal trichromats cannot do so. Each type of dichromat is characterized by a specific neutral band several nanometres in width (the mean values for the four dichromats are shown in Figure 6). Note that there are two such neutral points in the visible spectrum for the tetartanope but only one each for the other three dichromats.

The position of the neutral axis also denotes the broad *axes of confusion*, which for protanopes and deuteranopes is red-green, while it is yellow-blue for the tritanope and tetartanope. From the chromaticity diagram we can see that what distinguishes protans from deutans is not only the slight difference in the wavelength of their neutral points (493 nm for protanope and 497 nm for the deuteranope) but also the fact that protanopes confuse blue-green with *red* while the deuteranopes confuse blue-green with *purple*. It is such subtle differences, for example in reds, that must be employed in P.I.C. test plates if they are to distinguish between deutan and protan defects (that is, to be qualitatively diagnostic).

In the case of dichromats, the confusions are complete but this is not so for the extreme anomalous trichromats (E.P.A. or E.D.A.). In acquired dyschromatopsias involving the red-green axis, confusion loci are most often of the deutan type, while those for the yellow-blue axis resemble tritanopes more closely. If a line is traced perpendicular to the neutral axis for each defect, we will see which colours each dichromat is still likely to experience.

Combining this knowledge with colorimetric data obtained from colour vision tests, it is now possible to assess colour confusion theoretically from the chromaticity diagram, to predict whether it is intentional, and, if so, whether it is adequately realized. In tests implicitly designed to detect and diagnose dichromats, it is therefore possible to estimate whether the given design falls within the theoretical boundaries of maximal colour confusion according to the tenets of colour confusion theory.

Colour discrimination Perception of small colour differences in all parts of the colour space implies good discrimination, but even among normal

trichromatic observers there are individual differences, some possessing very 'acute' discrimination while others are so lacking in this quality that it amounts to partial confusion. The method of determining whether test elements lie on the dichromat's colour confusion loci does not provide adequate information about how difficult a test will be for those whose colour losses do not amount to outright dichromatism. A test situation that is difficult for dichromats is not necessarily so for those with minor colour defects, and an overall index of the colour differences found in the task provides us with a better measure. It is now possible to calculate these and also to scale the discrimination task in order of difficulty – the smaller the colour difference, the more difficult the discrimination task.

Such predictions from distances on the C.I.E. diagram cannot be made directly with any degree of accuracy or validity because the 1931 C.I.E. chromaticity space is not *visually uniform* and references to distances would not correspond with actual colour experience. It should also be remembered that no brightness differences can be inferred from such a diagram. However, when differences exist only because of hue or saturation, colour difference can be read off directly from a modified colour space, the so-called *uniform chromaticity scale* (or U.C.S. diagrams) first devised by Judd in 1935. But when there are also brightness differences, as is generally the case in most colour vision test tasks, colour differences have to be calculated from the original XYZ tristimulus values for the surfaces involved. There are a number of colour difference formulae, but the one used for calculations here is the Nickerson-Stultz formula (1942) in which the magnitude of colour difference, ΔC , is expressed as a unit bearing the name of the American National Bureau of Standards – the N.B.S. unit. One such N.B.S. unit is equivalent to a colour difference of about 5 just noticeable differences (j.n.d.) as perceived by a normal observer under ideal viewing conditions. To give an idea of what such a unit looks like, ΔC between the colours of the April and July issues of this journal has been calculated (for the actual tristimulus values refer back to Table 1). These differences are very small indeed but still measurable and ΔC between the two blue covers viewed under illuminant 'A' was found to be 1.5 N.B.S. units, whereas under illuminant 'C' it was slightly smaller, being 1.1 N.B.S. units. Although the colour differences quoted are minute, they help to illustrate the type of discrimination that would be expected of colour workers in industry. In the plastics industry, the tolerances for colour matches are even more critical and only differences of 0.5 of an N.B.S. unit or less are acceptable (that is differences involving only 2 to 3 j.n.d.). Such a task is very close to the best colour resolution that can be expected of

young people with perfect discrimination. In woollen textile industries higher tolerances are acceptable but even here the largest acceptable ΔC is only between 3 to 6 N.B.S. units. Thus, tests capable of selecting such colour workers must present an order of difficulty far greater than is usually found in tests for detecting colour confusions.

In the past, where individual stimuli might be only 5 to 10 mm. in diameter, it was difficult to make spectrophotometric and colorimetric measurements of surface colour. With certain modern instruments, however, reliable valid measurements can be made

(see Lakowski, 1964, 1965b, 1966) and therefore objective analysis is possible, and, by finding out what the colour differences in a test situation are, we are better able to choose the appropriate test capable of presenting close steps in discrimination. For further readings on the topic of colour specifications consult Le Grand (1968), Wright (1964), Stiles and Wyszecki (1967), and Judd and Wyszecki (1963).

The references for this paper will appear at the end of Part 2 in the October issue of the Journal.

Received for publication November 16, 1968.