

Using clinical tests of colour vision to predict the ability of colour vision deficient patients to name surface colours

Barry L. Cole¹, Ka-Yee Lian^{2,3} and Carol Lakkis^{1,2}

¹Department of Optometry and Vision Sciences, University of Melbourne, Victoria, Australia, ²Clinical Vision Research Australia, Cnr Cardigan and Keppel Streets, Carlton, Victoria, Australia, and

³Melbourne Optometry Clinic, Victorian College of Optometry, Melbourne, Victoria, Australia

Abstract

Purpose: To determine the predictive power of commonly used tests for abnormal colour vision to identify patients who can or cannot name surface colours without error.

Methods: The colour vision of 99 subjects with colour vision deficiency (CVD) was assessed using the Ishihara, the Richmond HRR (2002), the Farnsworth D15, the Medmont C100 and the Nagel anomaloscope. They named 10 surface colours (red, orange, brown, yellow, green, blue, purple, white, grey and black), which were presented in two shapes (lines and dots) and three sizes. The surface colours were also named by an age-matched group of 20 subjects with normal colour vision. The performance of the clinical tests to predict the CVD subjects who made no colour naming errors and those who made errors is expressed in terms of the predictive value of a pass $P_{(P)}$ and the predictive value of a fail $P_{(F)}$.

Results: The $P_{(P)}$ values of the tests were between 0.59 and 0.70 and $P_{(F)}$ values were between 0.77 and 1.00.

Conclusions: A 'mild' classification with the Richmond HRR test, especially if no more than two errors are made on the HRR diagnostic plates, identifies patients with abnormal colour vision who are able to name surface colour codes without error or only the occasional error. A pass of the Farnsworth D15 test identifies patients who will make no or few (up to 6%) errors with a 10 colour code, but who will be able to name the colours of a seven colour code that does not include orange, brown and purple. If protans are excluded, the predictive value for a pass $P_{(P)}$ for the Farnsworth D15 is improved from 0.59 to 0.70. The anomaloscope is not an especially good predictor of those who can recognise surface colour codes. However, an anomaloscope range >35 units identifies those who have difficulty in recognising surface colour codes, as does a fail at the Farnsworth D15 test.

Keywords: abnormal colour vision, anomaloscope, colour naming, Farnsworth D15 test, predictive value of colour vision tests, Richmond HRR test

Introduction

Surface colour codes have long been used to convey information and organise complex visual displays and are now used even more widely with the advent of

colour computer screens and inexpensive colour printing. Colour is a very effective coding dimension: when it is well used, it speeds information acquisition and reduces errors especially for complex visual displays (Christ, 1975; Luder and Barber, 1984; Macdonald and Cole, 1988). Colour also has a valuable role in communication, to identify objects, as in 'the man you want to see is wearing a red coat' or 'the suspect's car is green' or 'get the blue file for me'.

People with abnormal colour vision have a reduced ability to discriminate colour and many are often unable to differentiate the main colour categories. They may therefore be denied the functional benefits of colour

Received: 22 November 2006

Revised form: 18 January 2007, 27 February 2007

Accepted: 4 March 2007

Correspondence and reprint requests to: Barry L. Cole.

Tel.: +61 3 9347 7424; Fax: +61 3 9349 7498.

E-mail address: bcole@optometry.unimelb.edu.au

coding as colour codes use the main colour categories. There is good evidence that this is the case. Most observers with abnormal colour vision make errors naming surface colours (Cole and Orenstein, 2003; Cole *et al.*, 2006). They make errors naming the colours used to code the values of resistors and capacitors (Walraven and Leebeck, 1960; Voke, 1976) and identify electrical cables (Voke, 1976; Hovis *et al.*, 1994). They have problems naming the surface colours in computer displays (Bergman and Duijnhouwer, 1980; Verriest and Uvijls, 1980) including those used in rail (Ramaswamy and Hovis, 2004) and aviation (Mahon and Jacobs, 1991; Mertens and Milburn, 1996). They have been shown to be slower and make more errors extracting information from colour-coded flight instrument displays even when the colour is used redundantly (Cole and Macdonald, 1988).

Optometrists need to be able to advise their patients with abnormal colour vision about the extent of the difficulties they may experience with colour including advice on the difficulty they may have naming colours and recognising colour codes. This can be important even for very young patients because colour is used at school for identification and as a means of grouping objects or ideas. Snyder (1973) gives an account of the problems he had at school because of his inability to name colours.

At times optometrists may be required to give their opinion as to whether a person with abnormal colour vision is able to do a job in which colour coding is important. Elaborate colour coding is increasingly used in safety-critical operations such as computer displays for control of rail and air-traffic, of power and telecommunications networks and of industrial processes. To provide this advice to the colour vision deficient patients or their prospective employers, optometrists need to know the relationship between the results of clinical colour vision tests and ability to name colours.

There have been previous studies relating clinical colour vision tests to the ability to name colours, notably those of Walraven and Leebeck (1960), Voke (1976), Cole and Orenstein (2003), Ramaswamy and Hovis (2004) and Hovis *et al.* (1994), but we had the opportunity to add to the data available as part of an investigation of the effect of size and shape on the ability to name surface colours, which we have reported elsewhere (Cole *et al.*, 2006).

Method

The colour vision of 99 subjects with colour vision deficiency was tested using the Ishihara test (24 plate 1993 edition), the Richmond HRR test (2002 Edition, Richmond Products, Albuquerque, NM, USA), the Medmont C100 (Medmont Pty Ltd, Vermont, Vic.,

Australia), the Farnsworth D15 test (Richmond Products) and the Type 1 Nagel anomaloscope (Schmidt and Hensch, Berlin, Germany). The subjects were then asked to name 10 surface colours having the colours red, orange, brown, yellow, green, blue, purple, white, grey and black. The colours were presented in two shapes (dots and lines) and three sizes. A control group of 20 subjects with normal colour vision also named the surface colours.

The study protocol conforms to the provision of the declaration of Helsinki and was approved by the Joint Ethics Committee of the National Vision Research Institute, the Victorian College of Optometry and the Department of Optometry and Vision Sciences. All subjects, or their legal guardians in case of the minors, gave written informed consent to their participation.

Subjects

The subjects were 100 consecutively presenting male patients who had been referred for assessment to the colour vision clinic of the Victorian College of Optometry. One subject, aged 9, was excluded because he could not do the Nagel anomaloscope and was generally uncooperative.

The average age of the 99 subjects was 28 ± 11 years (range 8–52). All subjects had a visual acuity of at least 6/7.5 in the better eye and had no history of ocular disease. Subjects wore spectacles to do the tests if needed. The control group was 20 males (average age 35 ± 12 years; range 12–57) recruited from among routine patients and colleagues who passed the Ishihara test, had no signs of ocular disease and had a visual acuity (with spectacles if necessary) of at least 6/7.5 in the better eye.

Colour vision testing protocols

The Ishihara test and the Farnsworth D15 test were illuminated using a Macbeth easel lamp (illuminant C, 200 lux). The Richmond HRR (2002) test was given under fluorescent lighting (6300 K, 1280 lux, CRI 85).

The fail criterion for the Ishihara test was three or more errors. The fail criterion for the Farnsworth D15 test was two or more diametrical crossings in the arrangement of colours. The subjects with abnormal colour vision made two or more errors with the screening plates of the Richmond HRR and the severity of their colour vision deficiency was classified in accordance with the test's instructions. They were classified as 'mild' if they made errors only in the first five diagnostic plates, as 'medium' if errors were made in the next three plates, and as 'strong' if errors were made in the last two plates.

The anomaloscope was used to classify the subjects by the von Kries classification of abnormal colour vision.

Subjects made three matches of the red + green field to the yellow field and the matching range was then determined starting from the last of the three matches. The matching range was determined by the method of limits. The experimenter (K-YL) adjusted the red + green mixture control step by step and the subject used the yellow control to restore the match until the limit was reached at which the colour-match could not be restored. White light adaptation was used for subjects with moderate to severe colour vision deficiency.

The anomaloscope was not used for four subjects because of equipment malfunction. These subjects are presumed to be anomalous trichromats because three passed the Farnsworth D15 test and the other was classed as mild by the Richmond HRR test. They were classified as protan or deutan by the Medmont C100. The Medmont C100 test uses alternating red and green LED lights to measure relative red to green luminous sensitivity by flicker photometry and is highly reliable differentiating protan and deutan colour vision deficiency (CVD) (Metha and Vingrys, 1992). *Table 1* gives the number of subjects for each class of abnormal colour vision.

Colour naming task

The 10 surface colours were laser printed on six A4 size cards in two shapes (dots and lines) each in three different sizes, with a mid-grey background. Each card presented all 10 colours in one shape and one size. The order of the colours was different on each of the six sheets.

Table 1. Number of subjects in the study classified by type and severity of their colour vision deficiency where anomalous trichromats are classified as mild or moderate using the Farnsworth D15 test

	Number	Expected ^a
Protanomaly		
Moderate, pass D15 test	15	8
Moderate, fail D15 test	5	4
Protanopia	8	12.5
Total protan	28	24.5
Deuteranomaly		
Mild, pass D15 test	44	41
Mild, fail D15 test	16	21
Deuteranopia	11	12.5
Total deutan	71	74.5

^aThe expected numbers are based on the known prevalences in Caucasian societies in which protanopia, deuteranopia and protanomaly each have a prevalence of 1% among males, and deuteranomaly has a prevalence of 5% (Birch, 1993). About one-third of anomalous trichromats (Crone, 1961) and all dichromats fail the Farnsworth D15 test (Helve, 1972).

The dots had diameters of 17, 7 and 1.9 mm. The lines were 20 mm long and were 3.5, 1.9, and 1.0 mm wide. At the test distance of 400 mm, the dots subtended 2.4, 1.0 and 0.27° at the eye, and the width of the lines subtended 0.50, 0.27 and 0.14°. The length of lines subtended 2.9°. The 10 colours were in two horizontal rows of five colours on each card. The vertical distance between the centres of the rows and the horizontal distance between centres of adjoining dots or lines was 37 mm (5.3°).

The colours were chosen so that, in the judgement of the experimenters, each colour fell unequivocally within one of the 10 colour categories. The spectral reflectances of the coloured samples were measured by spectroradiometry and their chromaticities and luminous reflectance were calculated with respect to illuminant C. These are shown in *Table 2* and *Figure 1*.

The cards were placed one at a time on an angled stand in the light box, so that they were perpendicular to the line of sight of the subject and were viewed binocularly from a distance of 400 mm set by a fixed forehead rest. The cards were illuminated at 1280 lux by two GE Polylux 860 18 W tri-phosphor fluorescent lamps (manufacturer's specification: colour temperature 6300 K, colour rendering index 85) (GE, Fairfield, CT, USA).

Subjects were given the diagnostic tests and then asked to name the colours. The colour cards were presented in order of decreasing size with the dots first

Table 2. Chromaticity coordinates (x, y) and reflectance factors (Y%) of the 10 surface colours with respect to illuminant C (CIE 1931 colour space)

	x	y	Y%
Red	0.523	0.313	14.3
Orange	0.502	0.389	25.1
Brown	0.443	0.392	12.3
Yellow	0.415	0.466	65.0
Green	0.287	0.475	20.2
Background grey	0.303	0.305	53.1
Blue	0.172	0.183	11.3
Purple	0.278	0.217	10.7
White	0.299	0.299	72.0
Grey	0.303	0.305	22.2
Black	0.347	0.371	2.7

Note 1: The values of x, y and Y% are calculated with respect to the spectral energy distribution of illuminant C. The illuminant used in the experiment was not illuminant C but a fluorescent source with a correlated colour temperature of 6300 K so the values of x, y and Y% as observed by the experimental subjects will have been slightly different.

Note 2: Colorimetry was undertaken by a commercial laboratory accredited by the National Association for Testing Authorities using test method SJ&A 4.1.1 with 0°/45° viewing geometry and a spectroradiometer, a reference white tile and an incandescent calibrated source. The uncertainty for both x and y is ±0.001 and for Y% is 1.2.

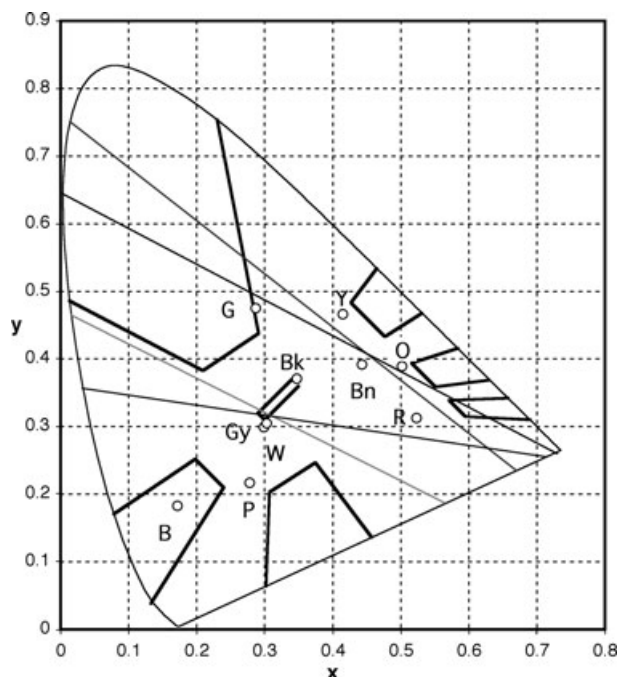


Figure 1. Commission Internationale de l'Éclairage (CIE) 1931 chromaticity diagram showing the chromaticities of the 10 colours named. The boxed areas define the domains of colours recommended by the CIE for surface colours used for signalling (CIE, 1983). The lines diagonally across the diagram are the confusion loci for protanopes (black lines) and deuteranopes (grey lines) that pass through illuminant A and illuminant C. These lines indicate the range of colours that look the same as daylight white (illuminant C) and warm white (illuminant A) for protanopes and deuteranopes when there are no brightness differences. All the colours lying in the area between the illuminant A and illuminant C confusion lines for each class of dichromat look white or grey to them.

or lines first for alternate subjects. There was one replication in which the cards were inverted, so the order of the colours was different. The order of dots and lines was reversed in the replication to balance practice and fatigue effects. The sizes were always presented in the same order from the largest to the smallest so that subjects had the benefit of practice with the larger sizes before naming the smallest size.

Subjects were told to use only the 10 colour names and they had in front of them a printed list of those names. Subjects were given up to 2 s (as judged by the experimenter) to name each colour. This length of time was rarely needed and naming of all 120 stimuli was usually completed within 4 min.

Results

The control group made no errors naming the surface colours but only 37% of those with abnormal colour vision made no errors. Errors made by the subjects with abnormal colour vision were dependent on the

type of the colour vision deficiency as determined by clinical diagnosis. The mean errors naming the surface colours for each type of abnormal colour vision are given in *Table 3*. It also shows the proportion of subjects in each type of colour vision deficiency who made no errors naming the colours. Full details of the errors made by the CVD observers and the effects of stimulus size and shape have been reported previously (Cole *et al.*, 2006). As reported in that previous paper, deuteranomals who passed the Farnsworth D15 test made significantly fewer errors than the other types of colour vision deficiency and protanomals who passed the Farnsworth D15 made significantly fewer errors than protanopes and deuteranopes. This means that standard clinical tests can be used in a general way to predict the ability of patients with abnormal colour vision to name surface colour codes. However, to be more precise in advising individual patients, it is useful to determine the predictive power of the clinical tests.

Predictive power of the clinical colour vision tests

The capacity of a clinical test of colour vision to identify those who can name surface colours as well as observers with normal colour vision can be expressed as (i) predictive value of passing, $P_{(P)}$, which expresses how well a pass at the clinical test identifies those who perform as well as those with normal colour vision and (ii) predictive value of failing, $P_{(F)}$, which expresses how well a fail at the clinical test identifies those who do not perform as well as those with normal colour vision. These terms are defined as:

$$P_{(P)} = \frac{\text{Number who pass the test and perform as well as those with normal colour vision}}{\text{Number who pass the test}}$$

$$P_{(F)} = \frac{\text{Number who fail the test and perform less well than those with normal colour vision}}{\text{Number who fail the test}}$$

Predictive value for a pass $P_{(P)}$ and predictive value for a fail $P_{(F)}$ should be distinguished from sensitivity and specificity. If the number of true positives = a , false positives = b , false negatives = c and true negatives = d , then $P_{(P)}$ is $d/(c + d)$ and $P_{(F)}$ is $a/(a + b)$. Sensitivity is $a/(a + c)$ and specificity is $d/(b + d)$. Predictive values are useful for clinicians because true and false positives (or negatives) are combined in one measure.

Table 4 gives the values of $P_{(P)}$ and $P_{(F)}$ for the Farnsworth D15 test, the Richmond HRR (2002) test and the anomaloscope range for various criteria for passing these tests. While none of the subjects with

Table 3. Mean per cent errors (all sizes and shapes) naming surface colours for each type of colour vision deficiency

Type of colour vision deficiency	Mean per cent errors	Per cent making zero errors
Deuteranomaly. Pass D15 test	0.7	70.5
Protanomaly. Pass D15 test	2.6	26.7
Deuteranomaly. Fail D15 test	4.4	12.5
Protanomaly. Fail D15 test	7.2	0
Deuteranopia	9.4	0
Protanopia ^a	11.0	0

^aExcluding one protanope outlier who made 42% errors naming surface colours.

normal colour vision made any colour naming errors, Table 4 also gives values of $P_{(P)}$ and $P_{(F)}$ if the limit of normal performance is defined as one colour naming error, which is an error rate of 0.8%. The values of $P_{(P)}$ and $P_{(F)}$ for various tests are similar and their confidence intervals overlap so there are no statistical differences between the tests. There is no significant difference between the first and second presentations in the number of errors made naming the colours.

Subjects made significantly fewer errors with the larger stimuli (Cole et al., 2006) and, in particular, the deuteranomals who passed the Farnsworth D15 test made very few errors. When only the large and medium-sized stimuli are considered, $P_{(P)}$ for a pass at the Farnsworth is 0.88 when one error is taken as the limit of normal performance. However, $P_{(F)}$ decreases to 0.72 because those who fail the Farnsworth D15 test also make fewer errors with the larger stimuli.

Table 3 shows that protanomals who pass the Farnsworth D15 test make more colour naming errors than deuteranomals who pass the D15 test and the difference is significant (Mann–Whitney U -test $p = 0.004$). It is therefore instructive to determine the predictive values of the Farnsworth D15 test for deuteranomals only. Table 4 shows the values of $P_{(P)}$ and $P_{(F)}$ for passing and failing the Farnsworth D15 test when protanomals are

identified and excluded by testing with the anomaloscope or the Medmont C100. When only the large and medium-sized stimuli are considered, 89% of the deuteranomals who passed the D15 test made no colour naming errors and 95% made either no errors or only one error.

It is particularly interesting to look at another combination of tests. Deuteranomals who passed the Farnsworth D15 test and were classified as ‘mild’ by the Richmond HRR made no more than one error. This equates to a $P_{(P)} = 1.00$. The $P_{(F)}$ for this combination of tests is 0.95, and the pass rate is 20%.

Figure 2 shows how the values of $P_{(P)}$ and $P_{(F)}$ and the pass rates change when the ‘fail’ criterion is based on varying number of errors made in the diagnostic plates of the Richmond HRR (2002) test. Figure 3 shows how the values of $P_{(P)}$ and $P_{(F)}$ and the pass rates change when different anomaloscope ranges are used as the ‘fail’ criterion.

Discussion

Identifying patients who can name surface colours

Table 4 shows that all the clinical tests are imperfect predictors of the ability to name surface colours, having predictive values of passing, $P_{(P)}$, of between 0.66 and 0.87 when one error of naming surface colours is taken as the limit of normal performance. There are no significant differences in the $P_{(P)}$ values of the various tests but taking into account the results of other studies, it is possible to comment on the relative value of the various tests for identifying CVD patients who are likely to make no or few errors with surface colour codes.

The Richmond HRR test may be the best test to identify those most likely to be able to name surface colours. In this study a classification as ‘mild’ found a $P_{(P)}$ value of 0.87, which means that 87% of the patients who are classified as ‘mild’ by the Richmond HRR test made no or very few ($\leq 0.8\%$) errors naming surface

Table 4. Predictive value of passing, $P_{(P)}$ and predictive value of failing, $P_{(F)}$ for naming of surface colours with zero errors or no more than one error^a by persons with abnormal colour vision and the percentage who pass the clinical test

Clinical test	Criterion for passing the clinical test	% passing the clinical test	Zero errors		No more than one error	
			$P_{(P)}$	$P_{(F)}$	$P_{(P)}$	$P_{(F)}$
Farnsworth D15	<2 diametrical crossings	60	0.59	0.95	0.73	0.90
	No error	44	0.61	0.82	0.75	0.75
Farnsworth D15 plus Medmont C100 or anomaloscope to exclude protans	<2 diametrical crossings with the Farnsworth D15 and deutan diagnosis with Medmont C100 or anomaloscope	44	0.70	0.93	0.84	0.85
Richmond HRR (2002)	Classification of severity as ‘mild’	30	0.70	0.77	0.87	0.70
Anomaloscope range	Range <30	67	0.64	1.00	0.66	0.97

^aAll subjects with normal colour vision make zero errors. The tolerance of one error is an error rate of 0.8%.

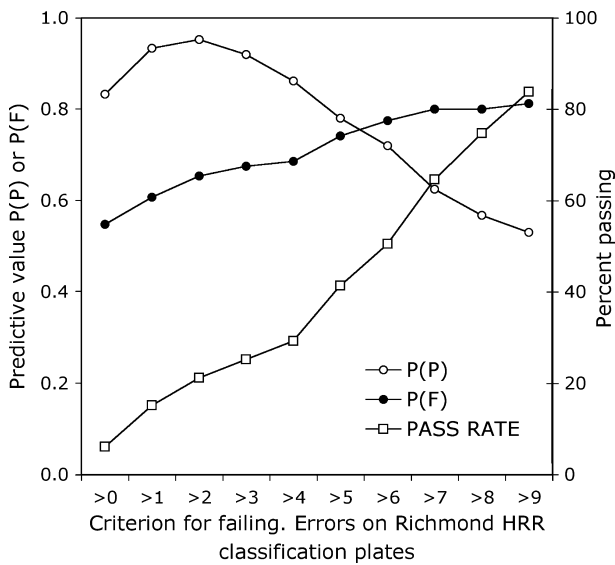


Figure 2. Predicted values of passing $P_{(P)}$ (open circles), predicted values of failing $P_{(F)}$ (filled circles) and pass rates (open squares) when the fail criterion is based on number of errors made on the Richmond HRR (2002) classification plates. The calculations are based on the assumption that one error naming the surface colours is the limit of normal performance.

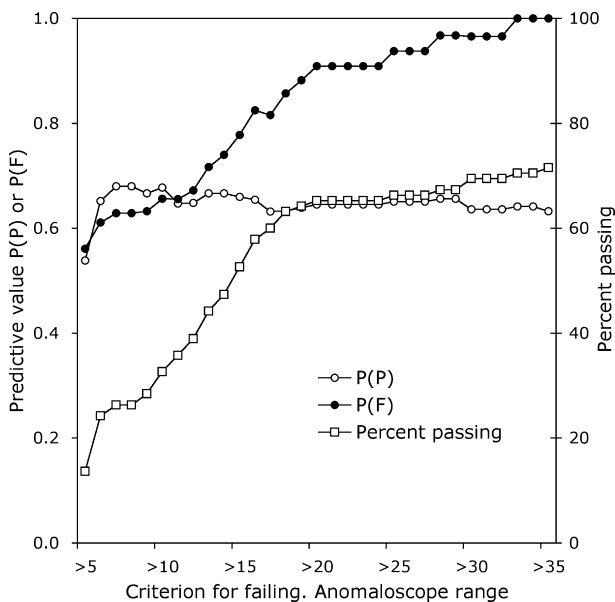


Figure 3. Predicted values of passing $P_{(P)}$ (open circles), predicted values of failing $P_{(F)}$ (filled circles) and pass rates (open squares) when the criterion for failing is based on the anomaloscope range. The calculations are based on the assumption that one error naming the surface colours is the limit of normal performance.

colour codes. *Figure 2* shows that $P_{(P)}$ for the Richmond HRR test can be as high as 0.95 if the criterion is no more than two errors on the HRR diagnostic plates, which means that almost all (95%) patients with defective colour vision who make no more than two

errors on the diagnostic plates will be able to name surface colour codes. This finding is supported by earlier reports. Ramaswamy and Hovis (2004) found $P_{(P)} = 1.00$ for the 1991 edition of the Richmond HRR test for predicting the ability of CVD subjects to recognise an eight colour VDU colour code when the criterion for passing the Richmond HRR test is no more than one error on the diagnostic plates. Walraven and Leebeck (1960) reported data showing that a ‘mild’ classification with the original version of the Richmond HRR test, the AO HRR, identifies observers who can name colour codes on resistors as well as normal observers ($P_{(P)} = 1.00$).

The Farnsworth D15 test is commonly used in clinical practice to dichotomise patients with abnormal colour vision as ‘mild’ or ‘severe’. Farnsworth (1943) designed the test so that those who passed it and are classified as ‘mild’ should be able to recognise colour codes used for electrical wires. Clinicians might well presume that a pass at the Farnsworth D15 test indicates the patient will be able to name surface colour codes. However, *Table 4* shows that the Farnsworth D15 test is an imperfect predictor of colour-code recognition. The same conclusion can be reached from other studies of surface colour naming by CVD observers. *Table 5* summarises the values of $P_{(P)}$ reported by or derived from other studies of surface colour naming from which it can be seen that the values of $P_{(P)}$ for the Farnsworth D15 test range from 0.69 to 0.82.

However, this does not mean that the Farnsworth D15 test has no value as a predictor of the ability to recognise surface colour codes. Cole *et al.* (2006) found that CVD subjects who pass the D15 test do not make many errors naming surface colours and mostly confuse red, orange and brown and blue and purple. They conclude that CVD observers who pass the Farnsworth D15 test can safely recognise the colours of a 7-colour code that omits orange, brown and purple, especially if the colours do not have too small an area.

Mild deuteranomals make significantly fewer errors naming surface colours than do mild protanomals (*Table 3*). Differentiating protans and deutans can help identify patients able to name surface colour codes: $P_{(P)}$ for the Farnsworth D15 test is 0.84 if protans are excluded (*Table 4*). Clinicians might protest that it is impractical to diagnose protan colour vision deficiency as few clinicians have access to an anomaloscope. However, the Medmont C100 test is a very reliable test for differentiating protans and deutans (Metha and Vingrys, 1992) and is inexpensive and easy to administer. It is a test that should be in every optometric practice as a part of the battery of tests for the assessment of colour vision.

The data from our investigation suggests that a combination of the Farnsworth D15, a test to identify

Table 5. Farnsworth D15 test predictive value of passing, $P_{(P)}$ and predictive value of failing, $P_{(F)}$ reported by or derived from various studies when the criterion for failing the Farnsworth D15 is two or more diametrical crossings

Study	Colour task and criterion for satisfactory performance	$P_{(P)}$	$P_{(F)}$
This study	Naming 10 surface colours of various sizes One error	0.73	0.90
Ramaswamy and Hovis (2004)	Naming 8 VDU generated colours 99th percentile of CVN observers >2 errors	0.70	0.90
Hovis <i>et al.</i> (1994)	Naming 10 colours of large gauge (3.5 mm) electrical wire colours 95th percentile of CVN observers	0.73	0.72
	Naming 10 colours of small gauge (1 mm) electrical wire colours 95th percentile of CVN observers	0.82	0.94
Cole and Orenstein (2003)	Naming 10 colours of paint, fabric and cotton samples: large sizes Two or more errors (max error made by CVN)	0.76	0.73
	Naming 10 colours of paint, fabric and cotton samples: small sizes Three or more errors (max error made by CVN)	0.69	0.80

CVN, colour vision normal.

and exclude protans, and a 'mild' classification with the Richmond HRR test could be a very good means of identifying those who can name surface colours. The values of $P_{(P)}$ and $P_{(F)}$ obtained with this combination of tests were 1.00 and 0.95 respectively.

The anomaloscope had a low $P_{(P)}$ of about 0.64 to 0.66 in this study (Table 4) which does not change much as the pass criterion is changed (Figure 3). This is interesting because many clinicians who have a special interest in colour vision and access to an anomaloscope place reliance on a narrow anomaloscope range as a trustworthy indicator of good colour discrimination. Patients with a narrow anomaloscope range are usually regarded as unlikely to have practical everyday problems with colour. For example the Joint Aviation Requirements in the European Union set an anomaloscope range of 4-scale units or less as indicating an ability to perceive the colours used in aviation (Squire *et al.*, 2005). Our data suggests that this might not be so for surface colour codes and the data of Squire *et al.* (2005) suggest it might also not be the case for signal light colours. Baraas *et al.* (2006) also found that the anomaloscope range fails to predict variations of colour constancy among anomalous trichromats.

However, Hovis *et al.* (1994) found quite different results. They found high values of $P_{(P)}$ for the task of naming the colours of thin electrical wires [$P_{(P)} = 1.0$ for an anomaloscope range < 10 for 1-mm wires and $P_{(P)} = 0.73$ for the 3.5-mm wires]. This may be due to the fact that the task of naming the colours of the wires was a difficult one even for those with normal colour vision. The 95th percentile of their colour vision normal subjects made about 20% errors for the 1-mm wires, which lowers the performance bar against which colour vision deficient subjects were judged. In our study, the colour vision normal control group was able to name the surface colours without error.

Identifying patients who cannot name surface colours without errors

Clinicians are on safer ground identifying CVD patients unable to name surface colour codes. The predictive values of failing $P_{(F)}$ for the various colour vision tests are high, especially for the Farnsworth D15 test and the anomaloscope, for which the predictive values for failing [$P_{(F)}$] are 1.0 or close to 1.0. Table 5 shows that other studies also found the Farnsworth D15 test gave high values of $P_{(F)}$. Hovis *et al.* (1994) found the anomaloscope range gave values of $P_{(F)}$ of 0.82–0.90 for naming the colours of wires.

This means that nearly all of those who fail the Farnsworth D15 test and all of those with an anomaloscope range of > 35 units perform less well than colour vision normal observers in naming surface colours. Optometrists can advise those who fail the Farnsworth D15 test or have an anomaloscope range > 35 units that it is almost certain they will make errors with surface colour codes and would not be suitable for employment in safety critical work that involved recognising colour codes. Even when the colour code is redundant, they can be expected, on the basis of the study of Cole and Macdonald (1988), to be slower and make more errors than those with normal colour vision.

Conclusions

Most patients who are classified as 'mild' by the Richmond HRR test will be able to name the colours of a 10-colour code, especially those who make no more than two errors on the classification plates. The majority (60%) of patients who pass the Farnsworth D15 will make no errors naming a 10-colour code and the remainder will make few errors. They may make up to

6% errors, mostly confusing red, orange and brown and very occasionally blue and purple, as is evident in Figure 8 of Cole *et al.* (2006). They will be able to name the colours of a 7-colour code that omits brown, orange and purple with no or very few errors. Mild deuteranomalies are less likely to make errors with surface colour codes than mild protanomalies. Thus most mild deuteranomalies who pass the Farnsworth D15 can be relied upon to recognise a 10-colour code and will be even safer with a 7-colour code. This will be especially so, if they are also classified as 'mild' by the Richmond HRR test.

A narrow matching range with the anomaloscope does not assure the ability to name surface colour codes but those with a matching range > 35 units will make errors. Likewise, patients who fail the Farnsworth D15 test are very likely to make errors with surface colour codes.

Acknowledgement

The authors are grateful to Dr Steven Jenkins of Steven Jenkins and Associates for making the colorimetric measurements.

References

- Baraas, R. C., Foster, D. H., Amano, K. and Nascimento, S. M. C. (2006) Anomalous trichromats' judgments of surface color in natural scenes under different daylight. *Vis. Neurosci.* **23**, 629–635.
- Bergman, H. and Duijnhouwer, F. (1980) Recognition of VDU presented colors by color defective observers. *Proc. Hum. Factors Soc.* **???**, 611–616.
- Birch, J. (1993) *Diagnosis of Defective Colour Vision*. Butterworth-Heinemann, Oxford.
- Christ, R. E. (1975) Review and analysis of color coding research for visual displays. *Hum. Factors* **17**, 542–570.
- Cole, B. L. and Macdonald, W. A. (1988) Defective colour vision can impede information acquisition from redundantly colour coded video displays. *Ophthalmic Physiol. Opt.* **8**, 198–221.
- Cole, B. L. and Orenstein, J. M. (2003) Does the Farnsworth D15 test predict the ability to name colours? *Clin. Exp. Optom.* **86**, 221–229.
- Cole, B. L., Lian, K.-Y., Sharpe, K. and Lakkis, C. (2006) Categorical colour naming by persons with abnormal colour vision. *Optom. Vis. Sci.* **83**, 879–886.
- Commission Internationale de l'Éclairage (1983) *Recommendations for Surface Colours for Visual Signalling*. Publication CIE 39.2. CIE, Paris.
- Crone, R. A. (1961) Quantitative diagnosis of defective colour vision. A comparative evaluation of the Ishihara test, the Farnsworth dichotomous test and the Hardy–Rand–Rittler polychromatic plates. *Am. J. Ophthalmol.* **51**, 298–305.
- Farnsworth, D. (1943) The Farnsworth–Munsell 100-hue and dichotomous tests for colour vision. *J. Opt. Soc. Am.* **33**, 568–578.
- Helve, J. A. (1972) A comparative study of several diagnostic tests of colour vision used for measuring types and degrees of congenital red–green defects. *Acta Ophthalmol. Suppl.* **115**, 3–64.
- Hovis, J. K., Lu, B. and Neumann, P. (1994) The ability of colour vision tests to predict performance in identification of wire colour. *Can. J. Optom.* **55**, 219–227.
- Luder, P. B. and Barber, P. J. (1984) Redundant color coding on airborne CRT displays. *Hum. Factors* **26**, 19–32.
- Macdonald, W. A. and Cole, B. L. (1988) Evaluating the role of colour in a flight information cockpit display. *Ergonomics* **31**, 13–37.
- Mahon, L. E. and Jacobs, R. J. (1991) Electronic flight information system displays and colour defective observers. *Clin. Exp. Optom.* **74**, 196–203.
- Mertens, H. W. and Milburn, N. J. (1996) Performance of color-dependent air traffic control task as a function of color vision deficiency. *Aviat. Space Environ. Med.* **67**, 919–927.
- Metha, A. and Vingrys, A. J. (1992) The C100: a new dichotomiser of colour vision defectives. *Clin. Exp. Optom.* **69**, 354–357.
- Ramaswamy, S. and Hovis, J. K. (2004) Ability of the D-15 panel tests and HRR pseudoisochromatic plates to predict performance in naming VDT colors. *Vis. Neurosci.* **21**, 455–460.
- Snyder, C. R. (1973) The psychological implications of being colour blind. *J. Spec. Ed.* **7**, 51–54.
- Squire, T. J., Rodriguez-Carmona, M., Evans, A. D. B. and Barbur, J. L. (2005) Color vision tests for aviation: comparison of the anomaloscope and three lantern types. *Aviat. Space Environ. Med.* **76**, 421–429.
- Verriest, G. and Uvijls, A. (1980) Colour vision defects and electronics. In: *Colour Vision Deficiencies V. Proc. 5th Symposium of the International Research Group on Colour Vision Deficiencies* (eds G. Verriest and E. B. Sreiff). Hilger, Bristol, pp. 346–350.
- Voke, J. (1976) The industrial consequences of deficiencies of colour vision. PhD thesis. The City University, London. Published in part as Voke J. (1978) Industrial requirements and testing of colour vision. *Mod Probl Ophthalmol* **19**, 82–87 and Voke J. (1980) Colour coding and the colour defective. *Ophthal Optician* **20**, 342–351.
- Walraven, P. L. and Leebeck, H. L. (1960) Recognition of color codes by normals and color defectives at several illumination levels. An evaluation of the HRR plates. *Am. J. Optom. Arch. Am. Acad. Optom.* **37**, 82–92.