British Medical Bulletin, 2017, 1–27
doi: 10.1093/bmb/ldx007

Invited Review

Colour vision requirements in visually demanding occupations

J. L. Barbur* and M. Rodriguez-Carmona

Applied Vision Research Centre, School of Health Sciences, City, University of London, Northampton Square, London EC1V 0HB, UK

*Correspondence address. Applied Vision Research Centre, School of Health Sciences, City, University of London, Northampton Square, London EC1V 0HB, UK. E-mail: j.l.barbur@city.ac.uk

Editorial Decision 13 January 2017; Accepted 20 February 2017

Abstract

Normal trichromatic colour vision (CV) is often required as a condition for employment in visually demanding occupations. If accurate CV tests were available to enforce this requirement, a significant percentage of subjects with anomalous, congenital trichromacy who can perform the suprathreshold, colour-related tasks encountered in many occupations with the same accuracy as normal trichromats would fail. These applicants would therefore be discriminated against unfairly. One solution to this problem is to produce minimum, justifiable CV requirements that are specific to each occupation. This has been done successfully for commercial aviation (i.e. the flight crew) and for Transport for London train drivers. An alternative approach is to make use of new findings and the statistical outcomes of past practices to produce graded, justifiable CV categories that can be enforced. To achieve this aim, we analysed colour assessment outcomes and quantified severity of CV loss in 1363 subjects. The severity of CV loss was measured in each subject and statistical, pass/fail outcomes established for each of the most commonly used, conventional colour assessment tests and protocols. This evidence and new findings that relate severity of loss to the effective use of colour signals in a number of tasks provide the basis for a new colour grading system based on six categories. A single colour assessment test is needed to establish the applicant's CV category which can range from 'supernormal', for the most stringent, colour-demanding tasks, to 'severe colour deficiency', when red/green CV is either absent or extremely weak.

Key words: colour vision, congenital colour deficiency, acquired colour deficiency, Ishihara plates test, Rayleigh match, CAD test

Background

Colour is by no means easy to define, but it is undoubtedly a perceptual attribute that affects significantly everything we see. A good place to start from when we question what we 'see' is the information available in the image formed on the retina by the optics of the eye. The ability to see spatially structured objects, to resolve edges and contours and to see fine detail requires point by point processing of the amount of light present in the retinal image and this is achieved through the photopic luminance contrast channel (see Fig. 1) which captures and averages the light present over the middle- and long-wavelength regions of the visual spectrum (Fig. 4). This channel acts as a single detector of light and does not therefore contribute to colour vision (CV). A second, colour-insensitive channel with similar properties, but with much coarser spatial resolution operates best at lower light levels and is based on spatial summation of rod signals. Variations in the spectral composition of light on the retina also carry useful information that is captured by the red/green (RG) and yellow/blue (YB) chromatic channels.

The perceptual representation of every object relies largely on the 'luminance' and 'colour' contrast signals the object generates in the eye with respect to its immediate surround. There are four principal 'Vision Information Channels' as shown in Figure 1. The contribution each of these channels makes to the perceptual representation of an object depends on a number of stimulus parameters such as the size of the object and its location in the visual field, the level of ambient illumination and the relative amounts of light reflected by the object in the short (S), middle (M) and long (L) wavelength regions of the visual spectrum. Although these 'channels' are a gross simplification of the mechanisms involved in the extraction of information from the retinal image, this simple model helps to explain some of the difficulties involved in colour assessment when the signals generated in RG and YB chromatic mechanisms have to be assessed independently.

What may not be immediately obvious is that the majority of colours we perceive as red, green, yellow or white and many shades of blue generate both RG and YB chromatic contrast signals and that the perceptual experience of what we see relies on the relative strengths of these colour signals and the luminance contrast of the object with some contribution from rods at lower light levels. It is therefore by no means easy to establish accurately how these signals interact and which channel contributes most to what we see in a given context.

Advantages of colour signals

In nature, colour signals carry important information that can often be used to great advantage.[1,2] When our surroundings are illuminated with ambient light, the reflecting properties of an illuminated object determine how much light is returned towards a viewing eye. The light that is returned is often altered in spectral composition by the spectrally selective absorbance of the object. This means that the returned light carries useful information about the chemical composition of the object. The human eye has evolved to make use of some of this information by comparing changes in the relative amounts of light that reach the retina in the short (S), middle (M) and long (L) wavelength regions of the visual spectrum. This comparison of signals (see Fig. 1) forms the basis for the RG and YB chromatic channels. The limits of the visual spectrum are determined by the spectral responsivity functions of S, M and L cones which extend from 400 to 800 nm (Fig. 1). Fruit, foliage and minerals often absorb light selectively within the spectral bands determined by the three cones. In the evolutionary context, this valuable information requires some form of easy to use perceptual representation. 'Colour' has evolved to enable this representation. It can therefore be defined as an important dimensional attribute of spatial vision that enables us to represent and use effortlessly information carried in the spectral composition of the light. In manmade environments, colours are often used to signal information which in turn can enhance many aspects of visual performance. Recent advances in colour display technologies have expanded greatly the number of applications that benefit from the use of colour. The following sections describe in greater

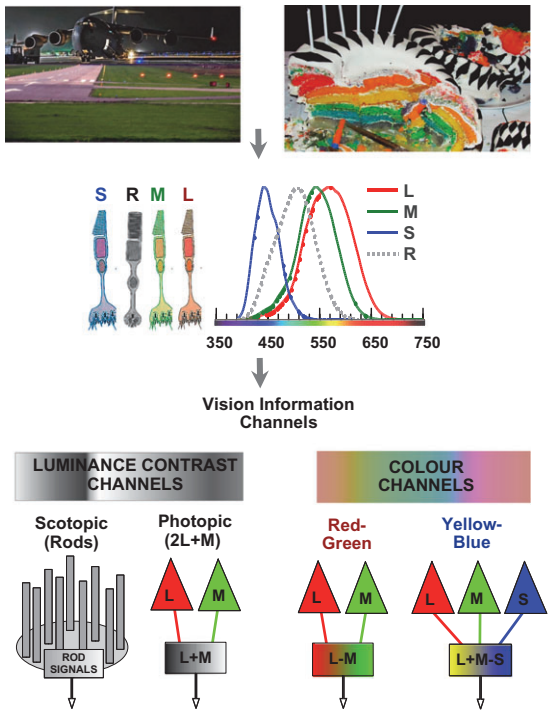

Fig. 1 Typical scenes where colour is useful, important and attractive. Information in complex scenes is extracted and channelled to the cortex efficiently via four principal 'vision information' channels. There are two achromatic channels which detect spatial differences in the amount of light in the retinal image over a large range of ambient illumination. These are often described as the 'photopic' and the 'scotopic' luminance channels. The photopic channel relies on the sum of L- and M-cone signals and carries no colour information. The scotopic channel relies on spatially pooled rod signals to achieve high sensitivity to light at the expense of spatial resolution. RG colour discrimination relies on differences in the amount of long-wavelength and middle-wavelength light as detected by L and M cones. S-cone signals reflect the amount of short-wavelength light present at a given point in the retinal image. Comparison of S-cone signals with the signals generated in L and M cones forms the basis for the YB channel. The signals generated in these two colour channels operate selectively through distinct mechanisms in the retina, but are then combined in later stages, when suprathreshold colour processing is involved, to yield the many colours we see.

detail how colour signals can enhance performance in visually demanding tasks.

Enhancement of object conspicuity

The 'conspicuity' of objects is largely a function of size and luminance contrast. Adding colour signals to an achromatic target defined by luminance contrast can enhance its conspicuity which means that the target is perceived to have higher 'effective' contrast. In the absence of colour, as in black and white displays, task completion times (TCTs) that involve visual search and other aspects of visual performance such as recognition of spatial cues and object properties depend largely on luminance contrast. When objects have low luminance contrast, the addition of colour signals, particularly to targets defined by luminance increments, results in improved visual performance and shorter TCTs.[3–5]

Parallel processing and pop-out

A fundamental property of colour mechanisms in human vision that yields significant advantages in visual search is the independent processing of colour and luminance contrast signals.[6–8] The existence of different visual mechanisms dedicated to the processing of colour signals also means that objects defined by colour are resilient to achromatic background clutter. Coloured targets are often picked up instantly and processed in parallel over large regions of the visual field. This phenomenon is called 'pop-out' and has been described in earlier studies.[9] Pop-out is especially useful when large, visually crowded displays are employed.[10] Because achromatic attributes are used in the form of text, shape, graphics and shading to represent detailed information, colour appears as a distinct dimension to create conspicuous differences between a target and distractors. Thus, pop-out of colour-coded information in complex scenes can be highly efficient and desirable.[9]

Segmentation and grouping operations

Colour can be used to group together subgroups of spatially discrete objects that are usually defined by luminance contrast. The human visual system organizes complex scenes into meaningful objects and/or spatially distinct regions. This is often described as 'segmentation'.[11] Visual segmentation can enhance performance and make the visual task less demanding and less tiresome. For example, an air traffic controller (ATC) can separate the aircraft on a

display into situation areas, or 'group together' the number of aircraft of immediate responsibility from the menu areas in a radar display. Thus, when controllers need to find a command in the menu bars, they can direct their attention to specific regions of the display. Segmentation tasks can be either regional or intended to group objects of interest into categories. Regional segmentation serves to separate a spatially continuous region from its surrounds, i.e. filling an area with colour to segment a restricted airspace from non-restricted airspace or using 'white' to group aircraft owned by a controller and 'green' to indicate un-owned aircraft makes the task easier to carry out. It has been shown that when this kind of regional segmentation is required, the use of colour is more effective than spatial cues defined only by luminance contrast.[12]

Signalling information by means of colour

The human sensory system can handle a large amount of information, but the bottleneck is often the cognitive system when parallel processing of information is not usually the norm. To speed up cognitive processing, one needs to reduce cognitive load by organizing complex information into categories defined by visual attributes that can be identified easily and often processed in parallel. Colour naming is a useful attribute one can use to convey information. Colours that are recognized easily can be linked directly to certain objects and can be given distinct meanings. Areas filled with red on a radar weather display are immediately associated with severe weather. In the ATC environment, identification of two stimuli is usually performed at separate spatial locations and can also be at different times. Typically, a controller remembers the colour by its name and searches for and identifies the target by its colour. Even in common activities, such as driving, colour is important and is often used to signal safety-critical information.[13] The use of red and white lights in Precision Approach Path Indicator (PAPI) lights in aviation is another good example. In this task, it is essential that the pilot can name correctly the number of red and white lights as an indication of the airplane's position with respect to the expected glide path designated for the runway. There are also other advantages. When certain objects are combined with specific colour signals, the objects become easier to remember in identification tasks when memory is required. Moreover, colour becomes increasingly more effective as an aid in the recall of memorized items.[14]

What happens in congenital deficiency?

Either the absence or the abnormal functioning of RG or YB chromatic mechanisms leads to reduced chromatic sensitivity (see Fig. 3b) as well as changes in the perceived colour of objects.[15] This in turn can cause a reduction in the 'effective' contrast or 'conspicuity' of objects with potential repercussions on visual performance. The RG channel compares signals from long-wavelength and middle-wavelength cones and contributes to RG CV, whilst S-cone output signals either yellow or blue when compared against steady L and M. In order for the YB colour channel to function properly, the retina must contain either functioning L or M cones or both, in addition to S cones. Normal trichromatic CV requires the normal functioning of all three cone classes. When one cone class is absent, the subject is a dichromat with only a functioning RG chromatic channel (in the absence of S cones) or only a functioning YB channel (in the absence of either L or M cones).

Anomalous trichromatic CV relies on the presence of three, spectrally distinct classes of cone pigments, but with at least one cone class shifted in its peak spectral responsivity (see Fig. 5a and b). When the retina contains normal S- and L-cone pigments and a variant L pigment (that substitutes functionally for the missing M), the person is described as being a deuteranomalous trichromat. When in addition to the normal S cone, the retina contains a normal M and a variant M pigment, the subject is described as being a protanomalous trichromat. The normal spectral separation between the peaks of L and M cones ($\delta\lambda_{max}$) is ~28 nm (with a midpoint at ~543 nm, see Figs 2b and 5c). The wavelength separation between a normal and a variant L

pigment can vary from ~1 to 12 nm (with a midpoint between the two peaks at ~551 nm), whilst that between a normal and a variant M pigment can vary from 1 to 7 nm (with a midpoint at ~532 nm, Fig. 2b).[16–18] Typical variant spectral responsivity functions for protanomalous and deuteranomalous trichromats are shown in Figure 5a and b. This means that when comparing the least affected congenital colour deficients, protanomalous subjects will be more disadvantaged than deuteranomalous trichromats (see Figs 2c and d, and 3a).

In terms of RG threshold sensitivity, congenital RG colour deficients always have higher thresholds than normal trichromats, but individual persons range from just above the normal upper threshold limits to almost complete loss of RG CV (Fig. 3). The gap between the least sensitive normal trichromat and the least affected deutan subject (see inset in Fig. 3a) is sufficiently large to separate all subjects with congenital colour deficiency (CCD) from normal trichromats. Figure 3b shows the extent to which the wavelength separation between L and M

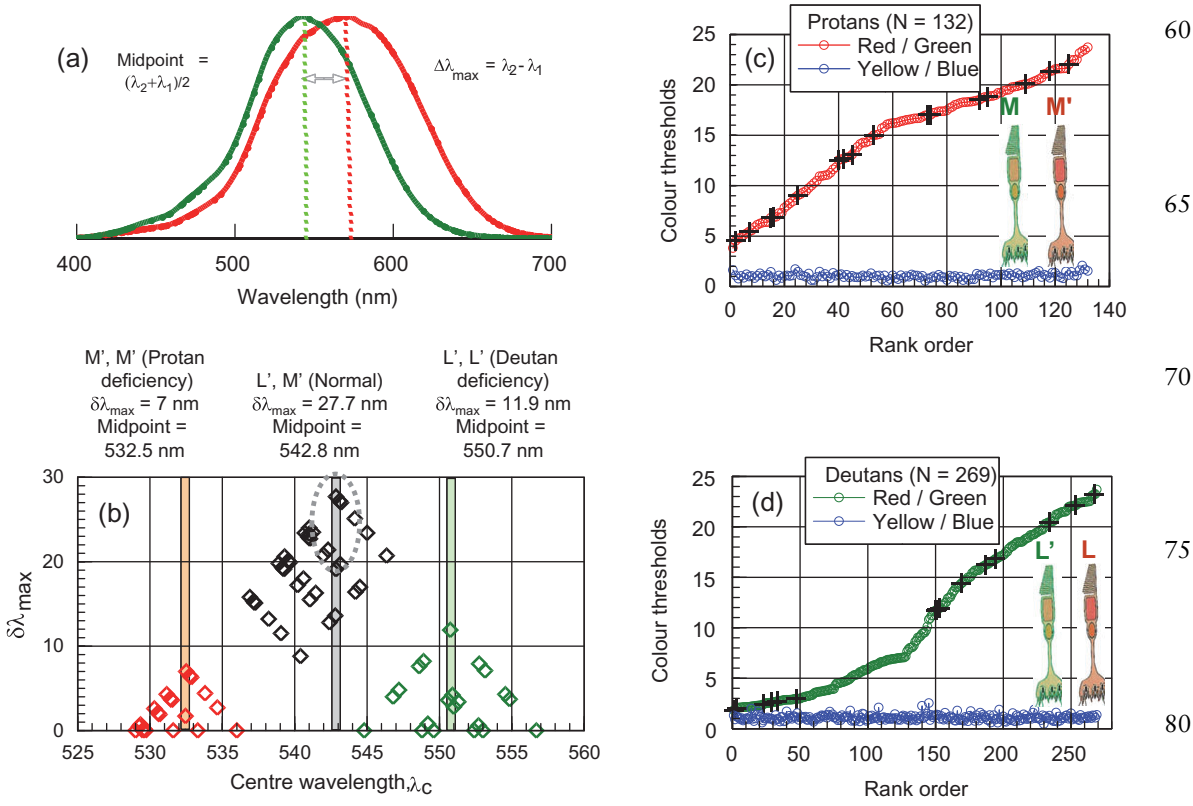

Fig. 2 The normalized spectral responsivity functions of L and M cones (**a**) reveal the wavelength separation, $\delta\lambda_{max}$, and the midpoint between the two peaks. Both parameters affect chromatic sensitivity and the outcome of colour assessment tests. The 'potential' pairs of L- and M-cone pigment variants one can associate with a normal trichromat (black symbols), a deutan (plotted as green) and a protan (red symbols) observer (**b**). Note that for the same wavelength separation, $\delta\lambda_{max}$, a protan and a deutan subject may perform quite differently in the same visual task. The RG and YB threshold colour signals measured in 132 protan subjects and ranked in increasing RG threshold are shown in (**c**). Note that a protan subject relies on a 'normal' M cone and a variant M′ cone pigment which substitutes for the missing L pigment. The opposite is the case for deutan subjects. The graphs (c and **d**) illustrate the severity of RG CV loss which ranges from close to normal thresholds to complete absence of CV. The colour thresholds are measured in standard normal (SN) Colour Assessment and Diagnosis (CAD) units[3] which are described and illustrated in Figure 6. One CAD unit represents the mean threshold colour signal measured in 330 young, healthy normal trichromats. The CAD units are shown as arrows in Figure 6a together with the age distribution of the subjects (Fig. 6f) and the spread in RG and YB thresholds measured in this group (Fig. 6d and e).

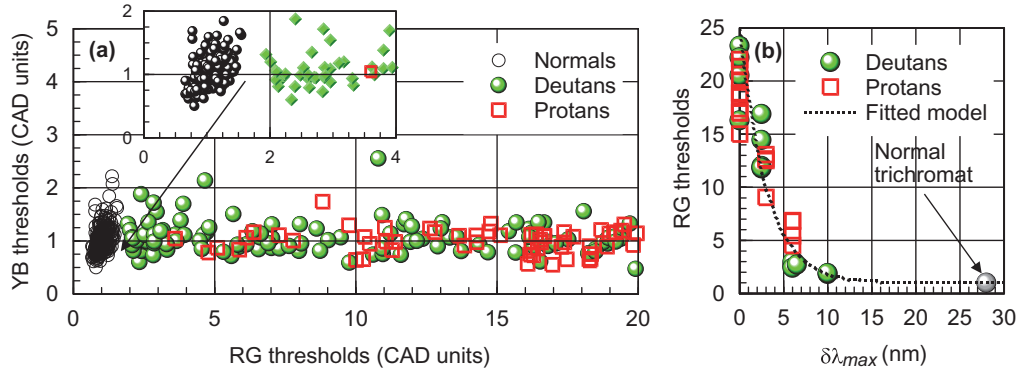

Fig. 3 YB thresholds plotted against the corresponding RG thresholds in 330 normal trichromats (black symbols), 269 deutans (green symbols) and in 132 protans (red symbols) (**a**). A subset of deutan and protan subjects had full genetic analysis of their cone pigment genes and corresponding, $\delta\lambda_{max}$ values estimated. The black crosses in Figure 2c and d indicate the subjects selected for genetic analysis. The RG thresholds for these subjects are plotted against the corresponding $\delta\lambda_{max}$ values in Section (**b**). Note that according to these results, wavelength separations greater than some 15–20 nm yield 'normal' RG colour thresholds[4] (reproduced and modified from Barbur *et al.*[22]).

cones affects the subject's RG threshold. The results show that normal levels of chromatic sensitivity can be achieved with $\delta\lambda_{max}$ values greater than ~18 nm.

Other parameters that affect RG chromatic sensitivity

The spectral separation between the peaks of L and M cones is not the only parameter that determines the strength of the signals generated in the RG chromatic channel and hence an individual's ability to carry out colour-related tasks. For the same spectral separation, the colour signals generated also depend on the 'mid-point' between the peak spectral sensitivities of L and M cones (see Fig. 2b). This effect contributes to increased inter-subject variability, particularly when CV is assessed on different tests.[19] In addition to these principal factors, there are also a number of other parameters that can affect the strength of RG colour signals generated when the eye is presented with what (to normal trichromats) are 'metamerically equivalent' RG colours. The latter appear perceptually identical, but can differ significantly in spectral composition. These parameters include the optical densities of L and M cones (which describe the amount of pigment present in each cone) as well as small shifts in the wavelengths of their peak spectral responsivity. 'Metamerism' in anomalous trichromats is not therefore equivalent to metamerism in normal trichromats. As a result, colour matches made by anomalous trichromats are not accepted by normal trichromats.[15] In addition, the relative numbers of L and M cones in the retina (i.e. the *L*/*M*-cone ratio)[20] can affect the overall RG sensitivity that can be achieved, particularly at low light levels, in the absence of background adaptation. Since the luminance contrast channel relies on the sum of signals generated in L and M cones, each of the parameters discussed above will have some effect on the luminous efficiency function, $V(\lambda)$, of the eye and hence on the luminance contrast of coloured stimuli[21] (see Fig. 4).

Achromatic stimuli that do not differ in relative spectral radiance from that of the immediate surround will always have the same luminance contrast, even for the extreme $V(\lambda)$ functions shown in Figure 4c that correspond only to either M or L cones. This is, however, not the case with coloured stimuli when the 'effective' luminance contrast can either be enhanced or diminished depending on the subject's class of colour deficiency and the colours of the objects and background involved. The parameters discussed will also affect the matching range and the midpoint of the RG mixture field when CV is assessed using the anomaloscope (see Fig. 12a) and this can lead to confusing results.[15,22,23] Differences in the optical density of cones does not

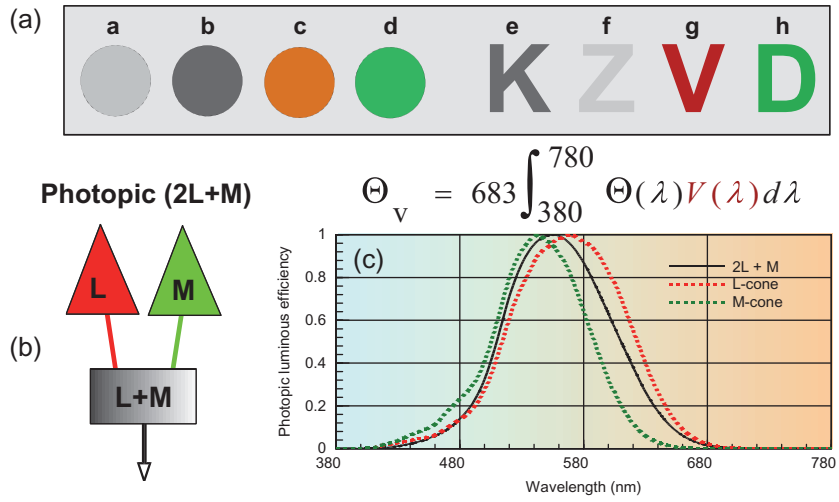

Fig. 4 The stimuli shown in (**a**) are either 'neutral' or are defined by both luminance and colour contrast. The spectral responsivity function, $V(\lambda)$, of the eye relies on summing up signals generated in L and M cones (**b**) and determines the retinal sensitivity to contrast. When the $V(\lambda)$ function receives only L or M cones (see dotted red and green curves in (**c**), respectively), the luminance contrast of objects defined by light of the same relative spectral composition as the surrounding background (i.e. the grey objects shown) remains unchanged. This is, however, not the case when coloured stimuli are involved. The $V(\lambda)$ function of the eye changes with *L/M* ratio in normal trichromats and is also strongly affected when variant pigments are involved (as in CCD). Coloured objects will therefore vary somewhat in luminance contrast in normal trichromats with much larger variations in subjects with colour deficiency. Protanopes and deuteranopes lack either L or M cones and hence experience the largest deviations in luminance contrast.

affect only their quantum catch (i.e. the number of photons absorbed per flash), but it can also change significantly the widths of their spectral responsivity functions with the largest width corresponding to the highest optical density. This means that some residual chromatic sensitivity is possible even when both L and M cones contain the same pigment.[24]

These realizations are important since they account for the significant variability which exists in normal trichromatic CV (see Fig. 6d and e). Equally importantly, they also account for the inconsistency of outcome when the same CCD subjects are assessed on different tests such as lanterns or tests that employ pseudoisochromatic plates with different illuminants.[19,25]

In spite of large variability that can be attributed, at least in part, to poor isolation of colour signals, differences in test protocol and inconsistent examination conditions, there is little doubt that the absence of RG CV can be a severe disadvantage in many occupational environments.[13] It is also of great interest to establish how the severity of loss in anomalous trichromats can affect visual performance in working environments that employ suprathreshold colours (often defined by both RG and YB colour signals). Studies carried out over several decades have shown that, in general, subjects with abnormal CV make more errors on many conventional colour assessment tests and produce slower motor responses than those with normal CV.[26] How important are these differences and how do they relate to the applicant's severity of CV loss?

In a study of visual displays used for the control of train movements,[27] the authors found that more than half of their 52 colour deficient observers could perform as well as the majority of colour normal controls when the task involved naming of eight suprathreshold colours generated on a visual display (red, yellow, green, blue-green, blue, purple, white and grey). Past studies that examined different

subject groups using different tests and visual tasks have not produced consistent findings. This is not too surprising since the severity of CV loss in subjects with congenital RG deficiency forms a continuum (Fig. 2c and d) from almost normal chromatic sensitivity to complete absence of CV.[18] The level of difficulty associated with colour-related tasks also varies considerably from task to task. In general, redundant cues are almost always involved and the colours used are always well above threshold and most likely to stimulate both RG and YB mechanisms. The latter is rarely affected by CCD. Nevertheless, even mild RG colour deficients confuse some colours, particularly those that rely mostly on small RG colour differences and this can be of concern in some applications.

Statistics of CCD

Subjects with normal trichromatic CV possess three distinct classes of cone photoreceptor, but significant variability exists even within the 'normal' group, see Figure 6d and e. Congenital RG colour deficiency affects ~8% of men and ~1% of women.[28,29] Only a small percentage of males exhibit complete absence of RG CV (Table 1).

Protanopes (~1%) lack functioning M cone and rely entirely on L cone for their spectral luminous efficiency function (Fig. 4c). The opposite is the case for deuteranopes (~1.1%). The cone sampling density remains unchanged and therefore visual acuity stays within normal limits. The absence or either L or M cones results in complete loss or RG CV, but protanopes and deuteranopes exhibit significant differences in their YB CV and have different luminous efficiency functions (Fig. 4c) with important functional consequences. Protanopes, for example, have much reduced sensitivity to long-wavelength lights (such as 'red' signal lights) which appear dark and less conspicuous.

Coloured stimuli also vary greatly in 'effective' luminous contrast when compared with normal trichromats and under rare circumstances can become completely undetectable by dichromats. On the other hand, coloured stimuli that are 'isoluminant' to a normal trichromat and are therefore only detected by means of colour contrast can have a large luminance contrast when seen by deuteranopes or protanopes and although they have little or no colour, they are easily detectable. Tritanopia implies the absence of functioning S cones and is extremely rare (~0.002%). If for any reason S-cone pigments were to shift in their peak spectral responsivity function (Fig. 5) in a way similar to the variant L and M pigments, this would have little or no effect on YB CV because of the large spectral separation between S and L or M cones. There are, therefore, no detectable congenital tritanomalous subjects. Anomalous trichromacy arises when a variant cone pigment substitutes for a missing 'normal' cone pigment. This results in subjects with either deuteranomalous (Fig. 5b) or protanomalous CV (Fig. 5a) many of whom retain some level of RG colour discrimination (Fig. 2b), but make abnormal metameric matches, as revealed in the anomaloscope test (Fig. 13).[22] The great majority of colour deficient subjects (~74%) are anomalous trichromats with severity of RG CV loss that forms a continuum with many exhibiting adequate residual CV to make them unaware of any loss (Fig. 3a and b). Although some of the traditional colour screening tests can be made sensitive enough to screen for normal trichromacy, this can only be achieved at the expense of specificity with many normal trichromats being classed as abnormal (see Table 2), an outcome that depends strongly on the number of errors allowed for a pass.

Table 1 The different classes and relative distribution of colour deficient subjects that make up ~8% of the male population[28]

Accepted prevalence of congenital CV deficiencies						
Protanope	Deuteranope	Tritanope	P-nomalous	D-nomalous	T-nomalous	Total
1	1.1	0.002	1	4.9	0	8.002

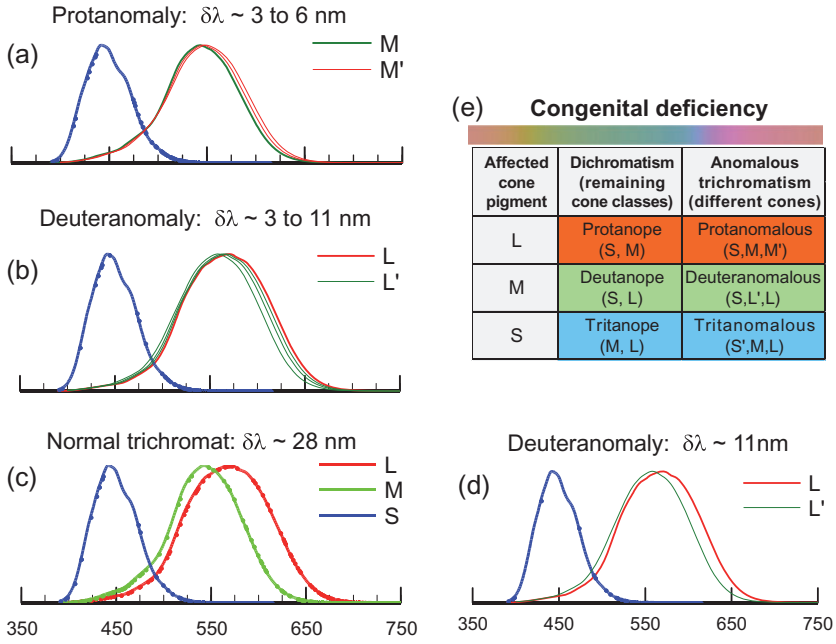

Fig. 5 Examples of CCD and the corresponding changes in the remaining cone pigments. The absence of S cones is rare (Table 1) and therefore YB CV remains normal in almost every subject with congenital RG deficiency. Dichromats lack one class of functioning cones and are classified as protanopes, deuteranopes or tritanopes (**e**). Anomalous trichromats differ from dichromats in that in addition to S cones and normal M or L cones they also have a third functioning cone which relies on either a variant M pigment (shown in red in Section **a**) or a variant L pigment (shown in green in Section (**b**)). These occurrences give rise to protanomaly (a) and deuteranomaly (b). The least affected deuteranomalous trichromat (**d**) has the largest separation between the normal and the variant L-cone pigments (~11 nm). The normal trichromat with a spectral separation of 28 nm is shown for comparison in (**c**).

An even greater problem is the difficulty of quantifying reliably the applicant's class of CV and the severity of loss.[19,30] The effects of monocular/binocular viewing, normal healthy aging and the loss of YB and RG chromatic sensitivity in acquired deficiency are also rarely considered.

Isolation of colour signals

In order to measure colour sensitivity, it is important to isolate the RG and the YB colour mechanisms by eliminating the detection of unwanted luminance contrast signals that are almost always present in coloured stimuli. Although many tests employ stimuli defined entirely by colour with no luminance contrast (and are often described as photopically 'isoluminant'), this is only true for the 'standard' human eye assumed to have normal L- and M-cone pigments and an *L/M*-cone ratio of 2.[31] Figure 4c shows the large variation in the spectral luminous efficiency function of the eye caused by variant pigment spectral responsivities and significant differences in *L/M* ratios in both normal and anomalous trichromats.[32] Luminance contrast masking is therefore employed in pseudoisochromatic plates in order to minimize the detection of residual luminance contrast signals that are also present in coloured stimuli.[33] Since it is important to isolate either RG or YB colour signals, the Ishihara (IH) test plates (Kanehara & Co. Ltd, Tokyo, Japan) also employ YB noise to isolate only the RG chromatic mechanism. This approach ensures that only subjects with good RG chromatic sensitivity pass the test. This high sensitivity for detection of congenital deficiency

Table 2 Summary of statistical classification outcomes in 1363 subjects based on the 38- and 24-plates edition of the Ishihara test

% classified as:	*N* (336)	*D* (705)	*P* (319)	*T* (3)
(a) Classification outcome based on zero errors (using the first 12 plates of the IH-24 plates edition)				
N	90.8	1.4	0.6	100.0
CVD	9.2	98.6	99.4	0.0
D	0.0	0.0	0.0	0.0
P	0.0	0.0	0.0	0.0
Indet	0.0	0.0	0.0	0.0
T	0.0	0.0	0.0	0.0
(Classification not possible: only 12 plates used)				
(b) Classification outcome based on 2 or less errors (using the first 12 plates of the IH-24 plates edition)				
N	99.7	6.7	1.6	100.0
CVD	0.3	93.3	98.4	0.0
D	0.0	0.0	0.0	0.0
P	0.0	0.0	0.0	0.0
Indet	0.0	0.0	0.0	0.0
T	0.0	0.0	0.0	0.0
(Classification not possible: only 12 plates used)				
(c) Classification outcome based on zero errors (using the first 25 plates of the IH-38 plates edition)				
N	81.85	0.71	0.63	100.0
CVD	18.15	99.29	99.37	0.0
D	0.00	60.14	8.20	0.0
P	0.00	0.14	38.49	0.0
Indet	18.15	39.71	53.31	0.0
T	0.0	0.0	0.0	0.0
(d) Classification outcome based on 3 or less errors (using the first 25 plates of the IH-38 plates edition)				
N	99.7	5.1	1.6	100.0
CVD	0.3	94.9	98.4	0.0
D	0.0	62.9	7.6	0.0
P	0.0	0.0	38.9	0.0
Indet	0.3	37.1	53.5	0.0
T	0.0	0.0	0.0	0.0

The results show the percentage of normal trichromats (N), deutan (D), protan (P) and tritan (T) subjects that are classified as normal, colour vision deficient (CVD), deutan, protan, tritan or indeterminate (Indet). Note that the statistical outcomes depend on the number of plates employed in the test and the number of errors allowed as a pass.

is, however, achieved at the expense of specificity. Although only 7 deutans and 6 protans in a thousand pass as normal when using the IH (38-plates edition) test with zero errors, as many as 18% of normal trichromats fail with at least one error.[34]

The Colour Assessment and Diagnosis test

Since luminance contrast signals are detected largely by transient mechanisms (i.e. magnocellular cells that respond best to rapidly changing stimuli), a more targeted way of isolating the use of colour signals is to bury the stimulus in dynamic luminance contrast noise[6,7] which masks luminance contrast signals. The Colour Assessment and Diagnosis (CAD) test is based on this experimental finding which has emerged from studies of camouflage. The technique provides effective masking of luminance contrast signals without affecting significantly chromatic sensitivity.[35] A colour-defined stimulus (see Fig. 6) is generated in a daylight (D_{65}) background and is presented to the subject moving diagonally in one of four possible directions. Following each presentation, the subject's task is to press one of four buttons (arranged to form a square) to indicate the direction of stimulus motion.[36] Figure 6a shows the background chromaticity at the centre of the ellipse and the chromatic displacement directions that isolate the YB and the RG chromatic mechanisms.

The cone contrasts generated by the moving coloured stimulus as a function of distance in the CIE (x, y) chart measured away from background chromaticity are shown in Figure 5b for the YB isolating axis and in Figure 5c for the RG axis. The results reveal two important properties of the CIE (x, y) 1931 chromaticity chart. First, chromatic displacements along the YB axis (Section a) isolate S cones without generating any signals in L and M cones (Fig. 6b). The RG axis, on the other hand, isolate L and M cones without generating signals in S cones (Fig. 6c). Second, the subject's thresholds for detection of either RG or YB colour differences are proportional to the cone contrasts generated by the stimulus. A doubling of thresholds therefore results in twice as large cone contrasts and hence chromatic saturation. This approximate linear relationship implies that the measured thresholds are proportional to the severity of CV loss.

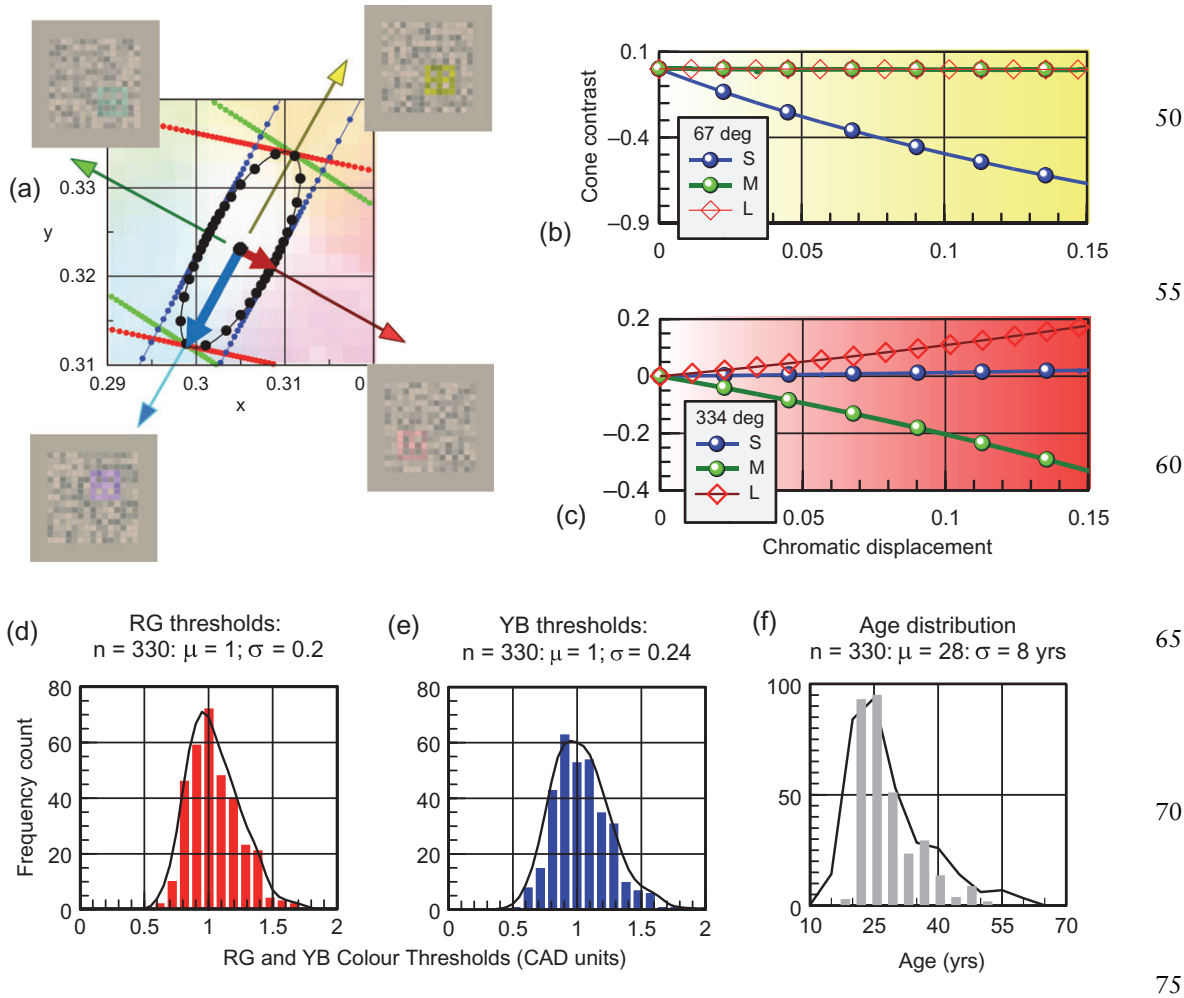

Fig. 6 RG and YB isolating axes are shown together with the threshold colour ellipse based on the mean RG and YB thresholds measured in 330, young, normal trichromats (**d**) and (**e**). The age distribution of the subjects is shown in Section (**f**). The minor and major axes of this ellipse define the SN CAD observer for RG and YB CV (see red and blue arrows in (**a**)). The screen dumps show the corresponding RG and YB isolating colours. Displacements away from background chromaticity along the YB and RG axes generate cone contrasts that are selective for S cones (**b**) and L and M cones (**c**). The larger these displacements, the more saturated the perceived colours. The measured thresholds are therefore directly proportional to the cone contrasts generated by the coloured stimulus.

The statistical distributions of RG and YB thresholds measured in 330 young, normal trichromats are shown in Figure 6d and e. The mean thresholds and the upper and lower statistical limits of these data define the standard normal (SN) CAD observer. The corresponding red and blue arrows in Section (a) indicate the mean thresholds for young normal trichromats. The measured thresholds are normalized with respect to the corresponding mean for each distribution. The results reveal a ~2.2-fold increase in threshold when comparing the least and most sensitive, young, normal trichromats.

Effects of normal aging

The variability shown in Figure 6d and e remains relatively constant in healthy normal aging, but the mean thresholds increase gradually with increasing age. Figure 7 shows RG (a) and YB (b) thresholds expressed in SN CAD units as a function of age.[37]

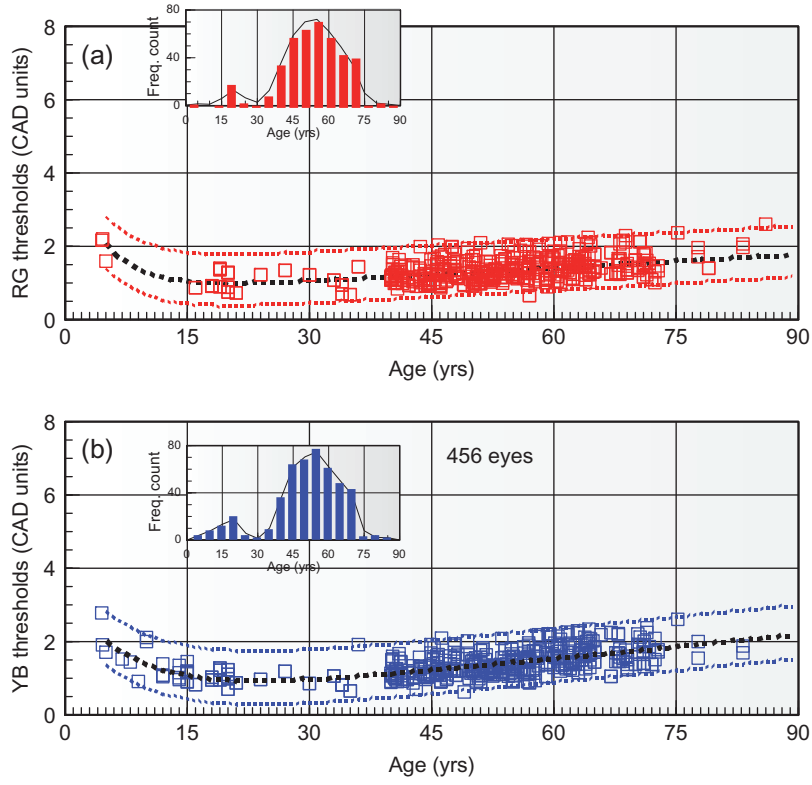

Fig. 7 RG (**a**) YB (**b**) binocular thresholds as a function of age for subjects deemed to have normal CV.[37] The histograms plot the age distribution of the study participants. The upper and lower limits represent ±2.5 SD. The mean binocular thresholds as a function of age are given by $RG_{bin} = 0.698 + 0.0121 \times age + 3.373 \times exp\ (-0.19 \times age)$, and $YB_{bin} = 0.24 + 0.0218 \times age + 2.99 \times exp\ (-0.1136 \times age)$. The smallest mean thresholds (~1 CAD unit) correspond to ~20 years of age. Above this age, mean thresholds increase linearly at a rate of ~1% per year for RG and ~1.6% for YB over the remaining life span (reproduced and modified from Barbur *et al.*[37]).

The results show that the best CV corresponds to ~20 years of age with a gradual, linear increase in thresholds of ~1% per year for RG and ~1.6% for YB, over the remaining life span. The mean and upper threshold limits as a function of age are important since they help define two categories of normal CV (see Discussion section).

RG loss of CV in congenital deficiency

Typical CAD results for a normal trichromat are shown in Figure 8a. The inner and outer ellipses plot the ±2.5σ limits for the corresponding age. The middle ellipse is based on the mean RG (minor semi-axis) and the YB thresholds (major semi-axis). The subject's colour thresholds are well within the normal range (RG = 1.22 and YB = 1.09). Deutan and protan-like subjects have normal YB thresholds, but larger RG thresholds (Sections b and c). The distribution of thresholds along the 12 RG directions employed in the CAD test allows for accurate, automatic classification of the class of colour deficiency involved. RG thresholds in CCD subjects vary from just above normal limits (Fig. 3a) to ~36 CAD units which is the limit imposed by the phosphors of the display employed (Fig. 8f). Subjects who hit the phosphor limits in the CAD test tend to also accept 'any' RG mixture as a match to the monochromatic yellow field in the anomaloscope test and are therefore classified as dichromats. The great majority of CCD subjects

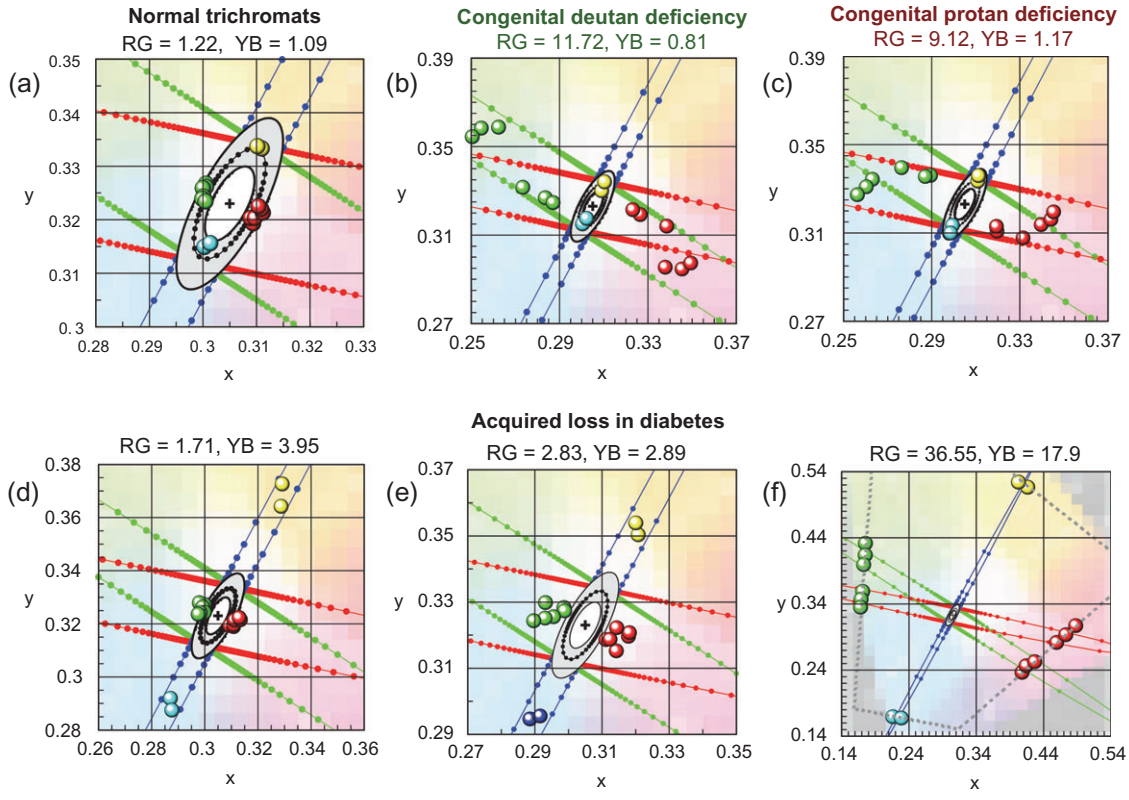

Fig. 8 Typical CAD thresholds are shown in (**a**) for a normal trichromat and in (**b**) and (**c**) for a subject with deutan- and protan-like congenital deficiency. The corresponding CAD thresholds for RG and YB CV are shown above each graph. Subjects with various degrees of loss of RG and YB chromatic sensitivity that are linked to diabetes are shown in (**d**), (**e**) and (**f**).

(~74%) are anomalous trichromats. A number of deutan and protan subjects (plotted as crosses in Figure 2c and d) had their cone pigment genes analysed genetically.[17] This information was then used to predict the spectral separations between the normal and variant pigments.[38] The results are plotted in Figure 3b and show that any spectral separation, $\delta\lambda_{max}$, between L and M cones greater than ~18 nm is likely to result in normal RG thresholds. Some subjects who, according to genetics, have only L- or M-cone pigments exhibit some residual RG chromatic sensitivity. This observation confirms earlier reports of residual chromatic sensitivity that can be linked directly to differences in pigment optical density, even when only one pigment class is involved.[39]

In summary, the severity of CV loss in deutan and protan subjects forms a continuum, as shown if Figure 2c and d. A number of discrete subgroups, determined largely by the $\delta\lambda_{max}$ values can be used to predict accurately the ranked distribution of RG thresholds in deutan and protan subjects.[18] This is possible simply because considerable variability in RG thresholds is present within each subgroup, largely as a result of differences in optical density, *L*/*M* ratio and possible differences in midpoint for the same $\delta\lambda_{max}$ (Fig. 2b). In conclusion, the continuous distribution of thresholds means that any categorization of CCD subjects has to rely on threshold limits with no clear-cut separation between categories in terms of severity of loss.

Loss of RG and YB CV in acquired deficiency

Acquired loss in patients with idiopathic disease, mostly over 55 years of age, is not uncommon with much greater losses in patients diagnosed with glaucoma,[40] age-related macular degeneration,[41]

optic neuritis and multiple sclerosis[42,43] or diabetes.[44] Both RG and YB chromatic mechanisms are affected in acquired loss, although during the earliest stages of disease, YB loss often precedes RG loss. Figure 8 shows CAD test results in three subjects with diabetes, but without retinopathy. Figure 8d shows a subject with acquired YB loss and relatively normal RG thresholds. Section (e) shows equal RG and YB loss, whilst Section (f), shows a diabetic subject with complete loss of both RG and YB CV. All three diabetic subjects had visual acuity within the normal range and no retinopathy.

Subjects with congenital RG colour deficiency can also develop acquired deficiencies. In general, RG thresholds that are significantly greater than YB thresholds, with the latter well above the upper, normal threshold limits, represents a clear sign of acquired loss on top of congenital RG colour deficiency. Since acquired loss of CV is not uncommon in older subjects, accurate assessment of RG and YB thresholds may well provide the earliest signs of retinal or systemic disease.

IH, Nagel anomaloscope and Holmes–Wright-Type A lantern tests

IH editions and standard industry protocols

The IH test is the most commonly used colour screening test within the electrical industry, primary healthcare, schools, universities and in many professional environments (see example plate in Fig. 9a). It is easy to administer and the interpretation of the results has an irresistible simplicity that often creates unforeseen problems. One might think it reasonable to assume that correct reading of all the plates with numerals is indicative of normal trichromatic CV, whilst errors made on every plate are equivalent to complete absence of RG CV. If this were the case, then the next logical step would be to make the subject's severity of CV loss proportional to the number of plates the subject fails, and this is what is often practiced. The numeral plates are not, however, equally difficult to pass and the probability of naming correctly a given plate depends on the subject's class of CV, i.e. whether the subject has normal, deutan or protan deficiency.[34]

The correlation between the number of plates the subject fails and the severity of RG CV loss is very poor. When the first 25 plates of the 38-plates edition are employed ~18% of normal trichromats make at least one error. The great advantage of using a classification outcome based on zero errors is that only 7 deutans and 6 protans in 1000 pass with zero errors.

The 38-plates edition when used with zero errors has high sensitivity and detects almost all CCD subjects. The problem is the large number of normal trichromats who make four or fewer errors. In addition, the few subjects who pass with zero errors (see Fig. 10a) are not always those with the least affected RG CV. One solution is the use of a secondary test,

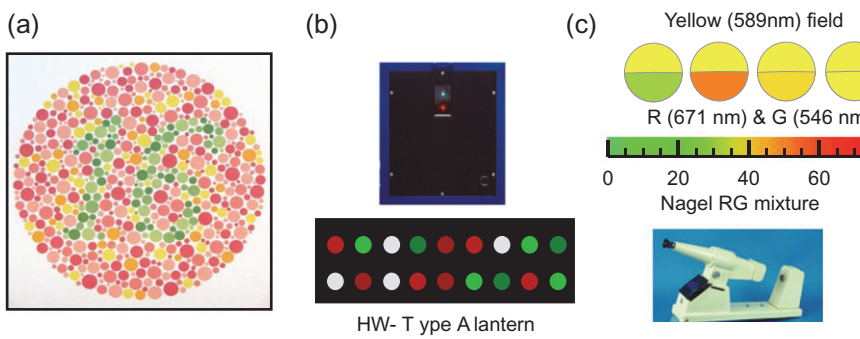

Fig. 9 Example of IH test plate (**a**), Holmes–Wright-A (HW-A) lantern and the colour pairs shown to the subject during the test (**b**) and the Nagel anomaloscope together with the RG mixture field and the monochromatic yellow field (**c**). To make a 'Nagel match', the subject has to adjust the relative amounts or 'red' and 'green' lights in the mixture field to match the colour of the monochromatic 'yellow' field.[22]

as was done by the UK Civil Aviation Authority (CAA) before the introduction of the CAD test in 2008. The CAA protocol employed the Holmes–Wright-Type A (HW-A) lantern as a secondary test. The HW-A is well designed with red, green and white lights presented as pairs (Fig. 9b). All normal trichromats pass the HW-A lantern using the internationally accepted protocol. Of note, 22% of deutan subjects also pass, but these deutans cannot be separated from the normal trichromats who pass, so they are therefore classed as normal. A second alternative is to use only the IH test, but to allow a certain number of errors as a pass. Colour assessment protocols based on three or fewer errors (Transport for London (TfL)) or eight or fewer errors (USA, Federal Aviation Administration) have been employed. Allowing 3 or more errors (on Plates 2–25) as a pass (see Table 2d) means that almost all normal trichromats pass, but 5.1% of deutans and 1.6% of protans also pass. Although the number of CCD subjects who pass is not large, some of these can have significant loss of RG CV (Fig. 10a). The great majority of those who pass will have thresholds below 5 CAD units.[34] To avoid these problems, TfL have developed a new protocol based on the CAD test and this has now been in use since 2008.[45]

It is of interest to establish the statistical outcome when the protocol employs the 24- or even the 14-plates edition of the IH test. The 24-plates edition is the most commonly used industry standard protocol based on errors made on the numeral Plates 2–13. This means that the subject can make a maximum of 12 errors. An analysis was carried out using 1360 subjects (705 deutans, 319 protans and 336 normal trichromats). The statistical outcomes of using this protocol are shown in Table 2b. Almost all normal trichromats pass with two or fewer errors together with 6.7% of deutan and 1.6% of protan subjects. The great majority of CCD subjects fail, but some of those who pass can have severe loss of RG CV with thresholds above 10 CAD units (Fig. 10b). The industry standard protocol based on the use of the IH-24 plates edition with two or fewer errors on Plates 2–13 is efficient, since fewer plates are involved, but the outcome is by no means equivalent to using the IH-38 plates system with two or even three errors allowed as a pass. When one continues to accept some errors, but some of the most challenging plates are eliminated, as is the case in the IH-24 and 14-plates edition, almost all normal trichromats pass, but the deutans who also pass with two or fewer errors tend to have more severe loss of RG CV because they fail less challenging plates (Fig. 10b). In addition, some protans with thresholds greater than six units also pass. These findings demonstrate clearly that by allowing a small number of errors and by using fewer plates one can

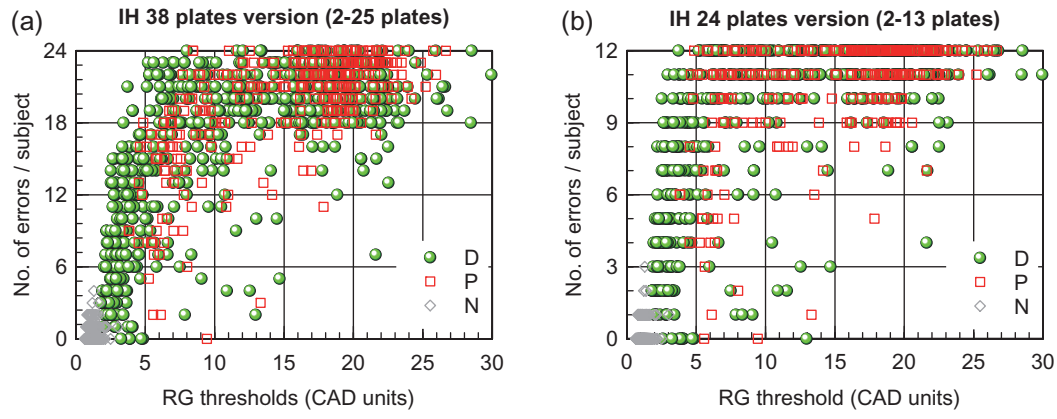

Fig. 10 The number of errors subjects make on Plates 2–25 of the IH-38 plates edition (**a**) and Plates 2–13 of the IH-24 plates edition (**b**). The latter is an efficient test because of the small number of plates involved and is used frequently in many industries with protocols based on two or fewer errors as a pass.

pass as 'safe' deutan and protan subjects with significant loss of RG chromatic sensitivity.

HW-Type A lantern outcomes

The HW-Type A lantern (see Fig. 9b) passes all normal trichromats when used with the CIE recommended protocol.[46] This lantern has been used as a secondary test by CAAs, the three services (army, air force and navy), the electrical industry, etc. HW-A and B lanterns are arguably the best designed and reliable lanterns with white, red and green colours of varying intensities that simulate signal lights. The statistical outcome of HW-A tests remains historically important since this lantern has been used successfully for decades in safety-critical occupations. In the absence of detailed studies designed to establish minimum CV requirements within a given occupation, it has become more important to establish pass/fail limits based on accurate assessment of the applicant's severity of CV loss that are at least equivalent to a pass on the HW-A lantern test. We have therefore examined the statistical outcome of the most commonly used protocol based on IH-24 plates, followed by HW-A as a secondary test (Table 3). This can be done by considering the statistical distribution of CCD subjects (Table 1) and the percentage of normal, deutan and protan subjects who pass the IH and HW-A tests (see Table 2b).

The advantages of using the HW-A lantern as a secondary test is that all the normal trichromats pass and that the number of deutans who pass increases from 6.7% to ~22%.

The introduction of the CAD test has replaced the need for a secondary test for pilots, ATCs, look out officers (within the maritime coastguard work) and other occupations that involve safety-critical and visually demanding tasks. In the absence of detailed studies designed to establish minimum CV requirements for specific occupational tasks, an acceptable alternative is to set pass/fail limits that are equivalent in outcome to those colour assessment protocols that have worked well in the past.

The HW-A lantern is not, however, without some problems. The lantern is no longer manufactured and there is no maintenance and calibration support. Repeated tests also exhibit some variability. The results shown in Figure 11b are based on the HW-A test results from 226 deutan subjects. The CIE recommended protocol was initially carried out in each subject to form two subgroups, those that passed (50 subjects) and those that failed (179 subjects). The probability of making, k, or fewer errors/run was calculated by repeating the

Table 3 Statistical outcome per thousand applicants based on a pass criterion of two or less errors using the IH-24 plates test followed by HW-A lantern test

Predicted outcome per 1000 applicants (IH-24 plates edition: 2 or fewer errors needed to pass on Plates 2–13)				
Applicants	1000	No. who fail IH	No. who fail HW-A	No. who pass
Normals	920	3	0	920
Deutans	60	56	47	13
Protans	20	20	20	0
Total	1000	78	67	933
% of applicants that undergo secondary tests			7.8	
% of normal who fail IH			0.33	
% of deutans who pass IH			6.7	
% of deutans who pass			21.7	
% of protans who pass			0.16	
% total colour deficient subjects who pass			16.6	

The calculations are based on the measured statistical outcomes shown in Table 2b.

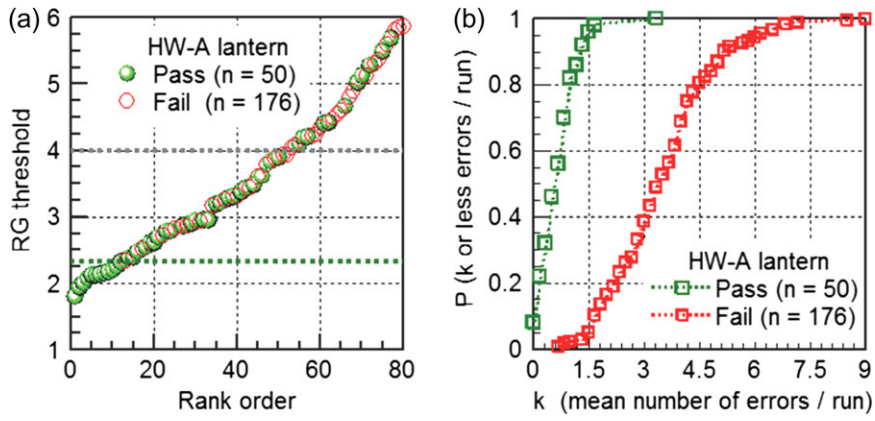

Fig. 11 Ranked distribution of CAD RG thresholds for the 226 deutan subjects (**a**) investigated. Section (**b**) shows a plot of the probability of making, *k*, or fewer errors per run for the deutan subjects that pass and for those that fail the HW-A protocol. Each subject carried out a total of 6 tests (i.e. 54 presentations of the paired vertical lights).

tests in the two subgroups 6 times (i.e. a total of 54 presentations of the HW-paired lights per subject). The results are plotted separately for those deutans that passed the HW-test protocol (green squares in Fig. 11b) and for those that failed (red squares). The results are important since they show clearly that even those who pass the first time can also make some errors in repeated runs and that several of the subjects who failed the HW-A protocol actually make fewer errors/run than some of the subjects who pass.

The results shown in Figure 11a suggest two useful RG threshold limits:

(a) All deutans with thresholds ≤2.35 SN CAD units (~just over 6% of the deutan population) pass the HW-A lantern protocol. This limit is shown by the dotted horizontal, green line in Figure 11a. This matches well the percentage of deutans who pass the IH-24 plates test with two or fewer errors using the industry standard protocol (see Table 2b). The only difference is that the 6% who pass have CAD thresholds ≤2.35 units and are therefore least affected in terms of their RG CV loss.

(b) About 22% of deutans who pass the HW-A lantern protocol have thresholds ≤4 SN CAD units. The number who pass with thresholds >4 CAD units equals the number who fail with thresholds ≤4 CAD units. A threshold ≤4 SN CAD units is therefore statistically equivalent to a HW-A pass.

Anomaloscope results and their interpretation

The relative amounts of red and green lights needed in the mixture field to match a monochromatic yellow field (Fig. 9c) are largely determined by the wavelengths of peak spectral responsivity of L and M cones with no contribution from S cones.

The exact amounts of red and green light in the mixture (i.e. the midpoint of the match) also reflect differences in the optical densities of the two pigments. The relative numbers of L and M cones in the retina affect the matching range which is often taken as a measure of RG chromatic sensitivity.[22,23] Subjects with CCD require either more red or more green light in the mixture field to match the colour appearance of the yellow field, depending on whether they rely on a variant L or M pigment to substitute for the missing, normal M or L pigment, respectively.[18] Figure 12 shows anomaloscope match midpoints ranked in increasing matching range for normal trichromats and for subjects with deutan and protan colour deficiency.

The anomaloscope is generally regarded as a useful and accurate instrument in distinguishing the protan from the deutan observer, and there is little doubt that the match midpoint and range reflect accurately

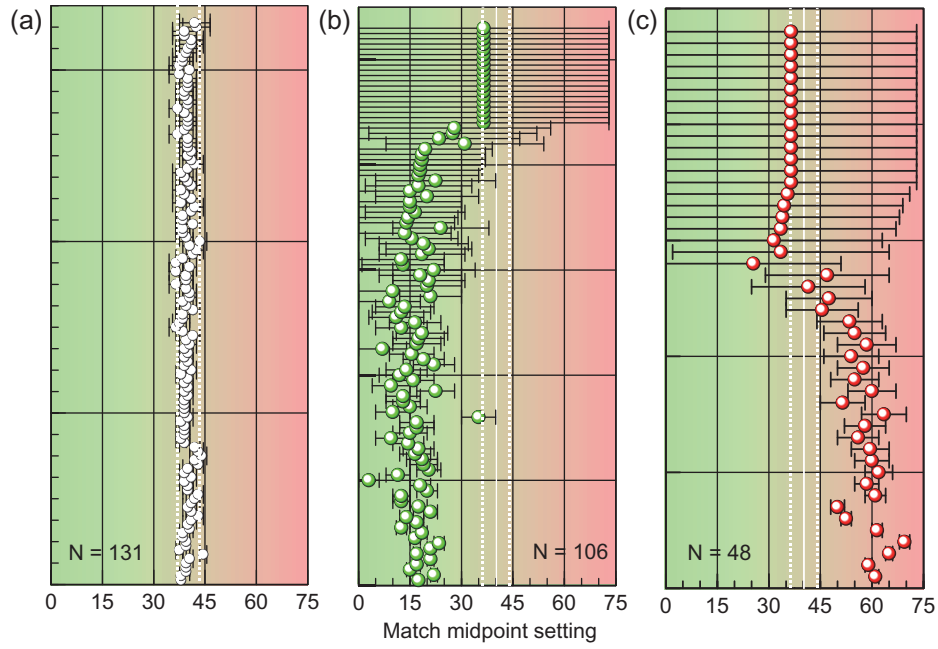

Fig. 12 The match midpoints are plotted in order of increasing matching range for normal trichromats (**a**) and for a number of deutan (**b**) and protan (**c**) subjects with varying severity of CV loss. The matching range is indicated by the horizontal bars that cross each midpoint. Dichromats can match the yellow field with any RG mixture. Some deutan and protan subjects exhibit very small matching ranges which suggests that although these subjects do not make normal matches, they are able to discriminate small changes in the RG mixture field as well as normal trichromats. (reproduced and modified from Barbur *et al.*[22]).

changes in the three principal parameters that affect the match. The variability in the measured anomaloscope match parameters is, however, large with matching range extremes of 1 and 9 units in normal trichromats. Some deutan and protan subjects exhibit very small matching ranges, often smaller than those measured in normal trichromats. Although the meaning of anomaloscope results is straight forward in the great majority of subjects, the match parameters are open to interpretation and do not always reflect correctly the subject's RG chromatic sensitivity. For example, some subjects produce Nagel matches with midpoints that fall just outside the normal range, but exhibit normal chromatic sensitivity, whilst a small number of subjects exhibit slightly reduced chromatic sensitivity, typical of minimal deuteranomaly, but make normal matches both in terms of midpoint and range (see Fig. 13).

In conclusion, anomaloscope matches can provide very useful information on the parameters that affect the match and in general distinguish accurately between deutan- and protan-like deficiencies. The parameters of the match cannot, however, be used reliably to quantify the subject's severity of CV loss.

Advantages of colour signals in large-field, visual search tasks

Applications designed for large-field, visual displays can make full use of colour to enhance many aspects of visual performance. The increased use of suprathreshold colours in many visually demanding occupations such as air traffic control makes it important to understand how diminished RG chromatic sensitivity can affect visual performance and also how to best design display applications to enable subjects with reduced chromatic sensitivity to perform as well as normal trichromats. In general, subjects with CCD tend to be slower and less accurate than normal trichromats in visual search tasks, particularly when low isolating RG colours and low chromatic saturations are employed.[47]

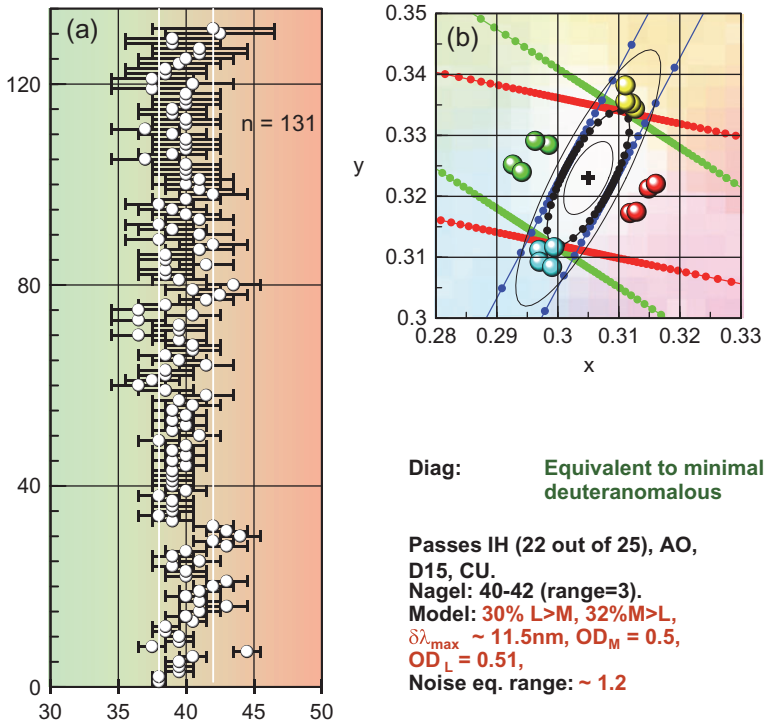

Fig. 13 The midpoints measured in normal trichromats and ranked in increasing matching range in Section (**a**) are the same as those plotted in Figure 11a, but on a magnified scale that shows the large variability which exists in normal trichromats. More importantly, the very few subjects that may rely on two variant genes make completely normal anomaloscope matches but exhibit slight loss of RG chromatic sensitivity, typical of the least affected deuteranopes. The subject shown in (**b**) makes completely normal Nagel anomaloscope matches (with a midpoint of 41 and a range of 3 units). The subject also passes D15, AO HRR and City University (2nd Ed.) CV tests and makes only two errors on the IH-38 plates edition.

The chromatic saturation employed in the visual search experiments described above was 10 CAD units for each of the four colours (a value that is within the range of chromatic saturations frequently employed in suprathreshold, real-life and colour-related tasks).

Figure 14 shows TCTs and percentage correct scores for a normal trichromat (with a RG CAD threshold of 0.7 units) and for the mild deuteranomalous subject (with a RG CAD threshold of 4.8 units). The task involved visual search when presented with multiple colours that isolate RG or YB chromatic mechanisms (a, c) or similar 'pastel' colours that always stimulate both RG and YB mechanisms. The RG isolating colours in the deutan subject yield longer TCTs and slightly reduced percentage correct scores (Fig. 14c). The same experiments were then repeated with the four 'pastel' colours, each of which generates both RG and YB chromatic signals (see screen dumps in Fig. 15). When presented with these colours, the deutan subject produced TCTs within the normal range and scored 100% correct responses in all trials. These and other experimental findings obtained in a study designed to investigate the use of colour in air traffic control applications (CAA (UK) report, CAP 1429 (2016)) demonstrate clearly that less severe deutans can perform visually demanding, colour-related tasks as well as normal trichromats provided saturated colours (i.e. >10 CAD threshold units) that generate both RG and YB chromatic signals are employed.

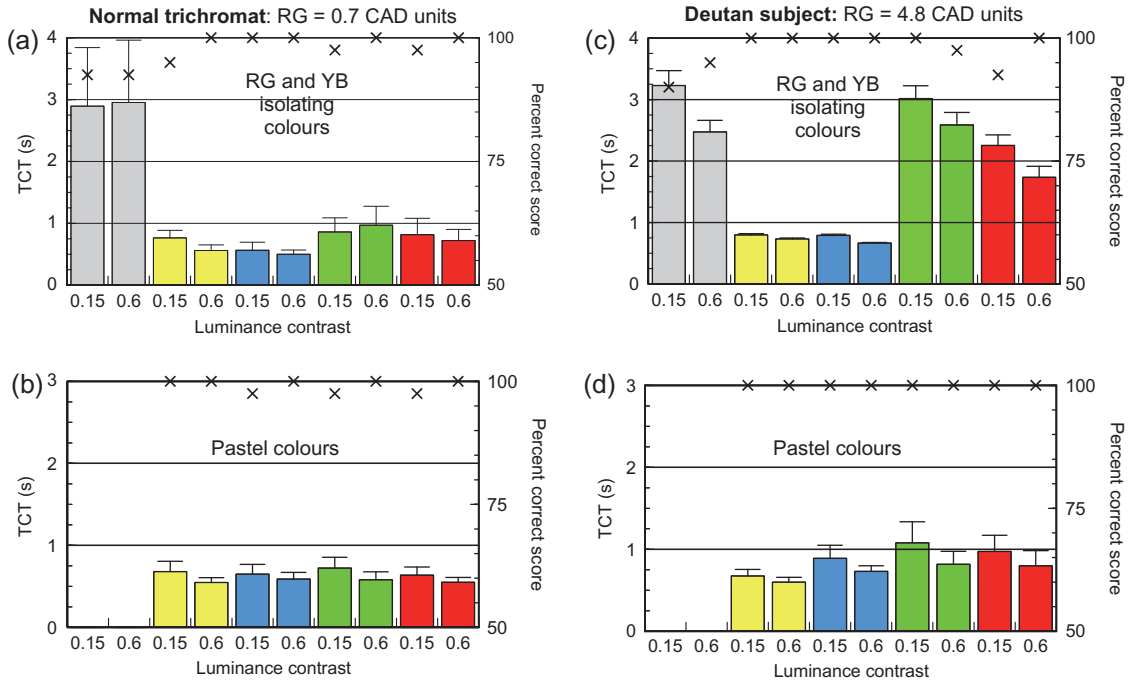

Fig. 14 TCTs with achromatic stimuli (grey bars) and when yellow, blue, green or red colours are added to the target. Data are shown for colours that isolate the RG and the YB chromatic mechanisms (**a**, **c**) and for 'pastel' colours that stimulate both RG and YB chromatic mechanisms (**b**, **d**). The deutan subject (with a RG threshold of 4.8 CAD units) requires longer TCTs for RG isolating stimuli, but performs as well as normal trichromats for the 'pastel' colours shown in Figure 15e. Percent correct scores are also displayed (black crosses) and reveal practically 100% correct scores in all the tests for both the normal trichromat and the deutan subject. Although the deutan subject requires longer TCTs when presented with RG isolating colours, he achieves 100% correct scores in all the tests.

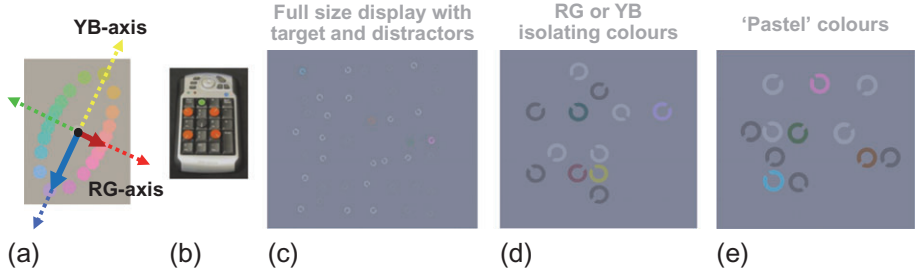

Fig. 15 Examples of visual search experiments involving coloured stimuli (**c**) and magnified regions to show the use of isolating RG and YB colours (**d**) and 'pastel' colours (**e**). The latter stimulate both the RG and the YB chromatic mechanisms and are therefore perceived to have some colour even by protanopes and deuteranopes.[3] The YB and RG colour isolating axes with respect to a D_{65} background are shown in (**a**) and the subject's response keypad in (**b**).

Discussion

The large sample data set examined in the main study reveals the statistical outcomes of the most commonly used colour assessment protocols. Particular emphasis has been placed on IH, Holmes–Wright lantern and CAD tests. The main study involved 1363 subjects with both normal trichromacy and congenital deficiency and the results were used to establish how CAD thresholds relate to the outcome of protocols that employ the IH test, often followed by the

HW-Type A lantern, as a secondary test. The results confirm the poor correlation between the number of errors subjects make on IH tests and the severity of CV loss. More importantly, the analysis carried out reveals significant differences in outcome between the various editions of the IH test. One important outcome of eliminating the most challenging plates in the 24- and 14-plates edition is to fail fewer normal trichromats. By using fewer plates, as in the 24-plates edition, and by allowing two or fewer errors as used routinely in industry standard protocols, almost all normal trichromats pass (Table 2), but some of the 6.7% deutans and 1.6% protans that also pass can have severe colour deficiency with thresholds as large as 13 SN CAD units (Fig. 10b). This is simply because when using the 24- or the 14-plates edition, the colour deficient subjects fail plates that are easier to pass. Although it is generally known that IH scores are not a good indicator of the severity of CV loss,[34,48] the new finding that subjects with severe loss of RG CV can pass the 24- or the 14-plates edition with two or fewer errors is particularly important. Even with three or fewer errors on the on the 38-plates edition, the very few deutans and protans who pass can have thresholds as large as 5–6 CAD units (Fig. 5).

Colour deficients have also been investigated on visually demanding, large-field, visual search tasks when speed of performance and response accuracy can be correlated to the subjects RG thresholds (CAA (UK) report, CAP 1429 (2016)). The findings from these studies, together with historical evidence of certification results based on conventional tests (i.e. IH pseudoisochromatic plates and the HW-A lantern tests) as well as current practices within occupations that rely entirely on IH-24 plates edition test have been examined with useful outcomes. When taken together these new findings justify the introduction of a new system of colour grading that can be enforced in practice, with categories that are statistically equivalent to the colour perception (CP) grades, currently used in the Defence Medical Services and in other environments. The CP grades rely heavily on the use of IH and the HW-A lantern protocols. The advantage of the new CV grading system is that the grades are based on accurate measurement of the applicant's RG and YB colour thresholds. The grades are therefore enforceable and this avoids the uncertainties involved in classifying the severity of CV loss (as shown in Figs 10–12).

In the absence of detailed studies designed to establish minimum CV requirements for specific occupational tasks (as has been done for flight crew in aviation and for TfL train drivers), an acceptable alternative is to consider carefully the new CV categories that have emerged from this study in view of the colour requirements in a specific occupation and to select the CV category that can be considered safe, without discriminating unfairly against those subjects with CCDs that can achieve levels of performance equivalent to normal trichromats.

A new CV grading system with application within occupations

'Normal' trichromatic CV (CV1)

This category includes all subjects with RG and YB CAD thresholds below the upper normal limits that have been established for healthy aging.[37]

'Functionally normal' trichromatic CV (CV2)

This category includes all applicants with a CAD threshold ≤2.35 CAD units. This limit is sufficient to pass all normal trichromats, irrespective of age and ~7% of the least affected deutans. The latter exhibit almost normal RG colour discrimination and 'pass' the HW-A lantern test 'with zero errors'. In terms of anomaloscope match parameters, the deutans who pass exhibit match ranges within normal limits, but require more 'green' in the RG mixture field to match the monochromatic yellow field. These subjects are not likely to have any colour detection and discrimination problems when suprathreshold colours defined by both RG and YB components are employed in visual displays.

'Safe' trichromatic CV (CV3)

This category includes all applicants with YB CAD thresholds within the normal range and RG thresholds ≤4 CAD units that cannot be classed as CV2. The higher limit is sufficient to pass all normal trichromats and ~22% of deutan subjects. This higher

limit matches the percentage of deutans who pass the HW-A lantern (22%) when using the CIE recommended protocol for use with this lantern. Although some of the deutans included in this group will have difficulties with small RG colour signals that are close to normal thresholds, all these subjects exhibit normal levels of visual performance when suprathreshold colours defined by both RG and YB components are employed in visual displays. In general, these subjects will not, however, accept metameric colour matches made by normal trichromats. The least affected protan-like subjects make errors on the HW-A lantern and exhibit minimum RG colour thresholds above 4 CAD units. As a result, few, if any, protan subjects can be included in this category. Only ~1% of protans and ~15% of deutans fall within the CV3 category (after removing those deutans that can be classed as CV2).

'Poor' RG CV (CV4)

This category includes all applicants with YB CAD thresholds within the normal range and RG thresholds ≤12 CAD units that cannot be classed as CV3. Subjects in this category have normal use of YB colour signals and can make some use of large RG colour signals. Working display environments often employ chromatic saturations as large as 24 CAD units, but only rarely as large as 36 CAD units. With chromatic signal strengths as large as 36 CAD units, the worst affected subjects within the CV4 category will still derive some benefit based on their RG CV. The benefit will, however, be small and comparable to what a normal trichromat can achieve with chromatic saturations as small as 3 CAD units. Based on our studies, ~32% of deutans and 29% of protans fall within this category. Many of them will have thresholds ~12 CAD units (range 4–12 units). With suprathreshold colour differences of ~24 CAD units, as frequently encountered in many display applications, the CV4 subjects benefit from significant use of colour with RG colour signals between 2 and 6 times above their corresponding threshold. Subjects in this category will therefore be able to make use of and cope with saturated RG colours on visual displays, but they will normally take longer to complete colour-related tasks and may also be less accurate (CAA (UK) report, CAP 1429 (2016)). When colour differences rely on both RG and YB colour signals and when the task employs large colour differences, such as the discrimination of reds and whites in the PAPI lights used in aviation, protan subjects with thresholds ≤12 CAD units perform as well as normal trichromats.[30] The same subjects can also carry out the less demanding, suprathreshold, colour-related tasks encountered in aviation. Since safety-critical task performance is not affected by their colour deficiency, protan subjects that fall within the CV4 category are allowed to work as commercial airline pilots.

'Severe' RG colour deficiency (CV5)

This category includes all applicants with YB CAD thresholds within the normal range and RG thresholds >12 CAD units. Approximately 70% of protans and 46% of deutan subjects fall into this category. Although some of the subjects included, i.e. those with thresholds just above 12 CAD units, can still make some use of saturated RG colours, the great majority have very little use of RG colour signals and have to rely mostly on YB colour differences. Many of these subjects will be unable to make much use of saturated RG colours in visual displays since the maximum chromatic saturations employed in many modern display applications rarely exceed 24 CAD units. There is also an additional disadvantage. Well-designed display applications employ both luminance and colour contrast to enhance visual performance. The use of adequate luminance contrast is essential when spatially structured patterns are involved and the visual task requires detection and discrimination of fine spatial detail. Luminance contrast is therefore an extremely important parameter, which together with stimulus colour and size, contributes significantly to the level of visual performance that can be achieved. In addition to very limited RG colour discrimination, subjects that fall into the CV5 category also experience significant changes in luminance contrast when viewing suprathreshold, coloured objects. Depending on the subject's class of colour deficiency and the

colours involved, the luminance contrast of an object can either be enhanced or diminished when compared with that perceived by normal trichromats. These changes can in turn affect significantly the visual performance they can achieve. Protan-like subjects that fall into the CV5 category, in particular, can also experience an additional disadvantage. When viewed against a bright background saturated, 'red' objects will appear highly conspicuous and 'dark' simply because they are often perceived to be of high negative contrast. When the same objects are viewed against a dark background, as is often the case with 'red' signal lights, they are much less conspicuous and can therefore go unnoticed. Good design of colour-related visual tasks, together with the use of 'pastel' colours defined by both RG and YB colour signals, in addition to luminance contrast, can minimize many of the disadvantages CV5 subjects can experience as a result of their CCD.

'Supernormal' trichromatic CV (CV0)

In addition to the five categories listed above, a new category, CV0, can be defined based on the availability of reliable data that describe variability in subjects with normal trichromatic CV.[37] This category includes only normal trichromats with RG colour thresholds below the mean value for the corresponding age. Only 50% of subjects with normal trichromatic CV can be included in this 'supernormal' category which may be useful when extremely demanding colour-related tasks are involved, particularly the naming of diffraction-limited signal lights under conditions of poor visibility.

CV grading categories for YB CV

Each of the six categories above assumes 'normal' YB CV. In practice, this is always the case when only congenital deficiencies are involved. Congenital tritanopia caused by the total absence of S-cone pigment is extremely rare (see Table 1). Congenital tritanomalous CV is also non-existent since any potential shifts in the wavelength of peak S-cone spectral responsivity, in a normally functioning YB system, either because of genetic factors or pre-receptor filtering of light in the eye, have a negligible effect on YB CV. This is simply because of the large wavelength separation which exists between S cones and M and L cones (Fig. 5). Congenital RG dichromats tend to have marginally smaller YB thresholds when compared with normal trichromats, and this can be attributed to the narrower wavelength sampling range of either the L or the M cones when compared with the opponent signal provided by the sum of L + M cones (in normal trichromats) (Table 4).

When the subject's thresholds exceed the upper limits that describe 'normal' YB CV as a function of age (Fig. 7b), the probability of acquired deficiency caused by diseases of the retina (such as glaucoma) or systemic diseases that also affect the visual pathways (such as diabetes) is high. Loss of YB CV, particularly when this loss is much greater than the corresponding RG loss, is highly indicative of retinal or systemic disease. RG thresholds significantly greater than YB thresholds with the latter well above the age-matched upper threshold limits observed in 'normal' vision are indicative of acquired loss in subjects that already have congenital RG deficiency. We have observed several such subjects in our studies which also exhibit another important characteristic that remains unexplained. In acquired colour deficiency, an increased stimulus size favours the YB threshold with reduced or no effect on RG threshold. These observations suggest that three categories are sufficient to describe YB CV.

'Supernormal' YB CV (CV0)

This category describes subjects with better than average YB CV for the corresponding age. This categorization can easily be established by examining the mean threshold limits that describe normal YB CV as a function of age.[37] RG dichromats tend to fall into this category because in general they have slightly smaller YB thresholds for the reasons discussed above.

'Normal' YB CV (CV1)

This category includes all subjects with YB CAD thresholds below the upper normal limits that have been established for healthy aging (Fig. 7b).

Table 4 Summary of the new CV categories that can be used to classify CV in normal trichromats and in subjects with congenital or/and acquired deficiency

The CV# grading system			
RG category	RG threshold requirement	Name	Description
CV0	$\leq RG_{mean}$	'Supernormal' CV	Includes only normal trichromats with RG thresholds ≤ the mean threshold for the corresponding age
CV1	$\leq RG_{upper}$	'Normal' trichromatic CV	Includes all normal trichromats (RG thresholds ≤ upper normal threshold limit for the corresponding age)
CV2	RG ≤ 2.35	'Functionally normal' CV	Includes all subjects with RG thresholds ≤2.35 CAD units that cannot be classed as CV1
CV3	RG ≤ 4	'Safe' CV	Includes all subjects with RG thresholds ≤4 CAD units that cannot be classed as CV2
CV4	RG ≤ 12	'Poor' RG CV	Includes all subjects with RG thresholds ≤12 CAD units that cannot be classed as CV3
CV5	RG > 12	'Severe' RG colour deficiency	Includes all subjects with RG thresholds >12 CAD units
YB category	YB threshold requirement	Name	Description
CV0	$\leq YB_{mean}$	'Supernormal' YB CV	Includes only subjects with YB thresholds ≤ the mean threshold for the corresponding age
CV1	$\leq YB_{upper}$	'Normal' YB CV	Includes only subjects with YB thresholds ≤ upper normal limits for the corresponding age
ACD	$>YB_{upper}$	Acquired colour deficiency	Includes all subjects with YB thresholds > upper normal CAD limits for the corresponding age

These CV grades include the principal categories of the CP grading system used by the armed services which relies heavily on IH and the HW-A lantern protocols. The advantage of the new CV grades is that they are based on accurate measurement of the applicant's RG and YB colour thresholds. The grades are therefore enforceable and avoid current uncertainties involved in classifying the severity of CV loss (see Table 2 and Fig. 10).

'Acquired' colour deficiency

YB thresholds above the upper normal, age-matched limits together with RG thresholds that are also frequently above the corresponding threshold limits that describe normal aging provide reliable indicators of acquired loss of CV and hence the presence of disease. Although YB colour sensitivity is often affected first in age-related macular degeneration, optic neuritis and diabetes, RG CV losses soon follow. Greater loss of RG CV (when compared with YB loss) is not normally observed in acquired deficiency, except in subjects with congenital RG deficiency.

Conclusions

Conventional colour assessment tests such as IH, lantern tests and the anomaloscope can be used effectively with appropriate protocols to screen for congenital deficiency with high sensitivity. These tests provide poor estimates of the severity of CV loss and when used in this way many normal trichromats also fail, which results in low specificity. More importantly, the use of protocols that allow for several errors and employ versions of the IH test with fewer plates (i.e. the 24- and 14-plates edition) fail only a small percentage of normal trichromats, but the deutans and protans who also pass can have severe loss of RG CV.

There is little doubt that key aspects of visual performance can be enhanced by adding RG and/or YB colour signals to targets and areas already defined by luminance contrast. Subjects with congenital RG colour deficiency will benefit less than normal trichromats from the addition of RG colour

signals. Although in some subjects with congenital deficiency the benefit remains substantial and can lead to normal levels of visual performance in suprathreshold, colour-related tasks, in general the benefit that can be achieved will be determined by the subject's severity of CV loss. CV requirements vary in employment from very demanding to non-essential. Detailed studies should ideally be designed and carried out to establish minimum CV requirements for specific occupational tasks (as has been done for flight crew in aviation and for TfL train drivers). In the absence of such studies, an acceptable alternative would be to consider carefully the six categories that have emerged from recent studies and also the analysis of pass/fail outcomes that emerge from the most commonly used conventional, colour assessment tests and protocols.

By examining specific CV requirements identified as important within a given occupation, one can then select one of the six categories that can be considered adequate and safe. Each of the specified CV categories can be enforced since the protocol relies only on accurate assessment of RG and YB colour thresholds under conditions that isolate the use of RG and YB colour signals and in the absence of other cues.

This approach can minimize unfair discrimination against those subjects with CCDs that can achieve levels of performance equivalent to normal trichromats. As an example of the benefits involved when adopting an appropriate colour grading category, consider the current, very common, industry standard protocol based on the use of Plates 2–13 of the IH-24 plates test with two or less errors as a pass. The outcome (as shown in Table 2 and Fig. 10b) is that 6.7% of deutans and 1.6% of protans pass, but some of those who pass can have severe loss of CV with RG thresholds as large as 12 CAD units. CV3 may well be the most appropriate category to replace this current practice. The adoption of this category would ensure that 22% of deutans and ~1% of protans pass, but more importantly, the severity of colour deficiency of those who pass would be reduced 3-fold to ~4 CAD units (Fig. 10b).

Acknowledgements

We wish to acknowledge the UK Civil Aviation Authority, Transport for London and the COLT foundation for supporting this work. Finally, we wish to thank the many subjects who participated in the numerous experimental studies carried out as part of this project.

Authors' biography

Professor John Barbur developed the Colour Assessment and Diagnosis Test (CAD) and other Advanced Vision and Optometric tests (AVOT). The intellectual property rights for the CAD test are owned by City, University of London and the UK Civil Aviation Authority. The test is manufactured and supplied by City Occupational Ltd, a spin out company of City, University of London. John Barbur is a director of City Occupational Ltd. and benefits (together with two other inventors) from the sale of AVOT tests.

Dr Marisa Rodriguez-Carmona studied Physics at Imperial College London where she graduated in 2001. She then joined the Applied Vision Research Centre at City, University of London as a doctoral student to work on 'Variability in chromatic sensitivity'. She received her doctoral degree in 2006 and continued as a Research Fellow working on various aspects of colour vision until 2015 when she became a lecturer in Optics and Visual Science. Dr Rodriguez-Carmona is not involved with City Occupational Ltd. and receives no financial benefits from the sale of AVOT tests.

Conflict of interest

None declared.

References

1. Martin DR, Fowlkes CC, Malik J. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Trans Pattern Anal Mach Intell*. 2004;26:530–49. doi:10.1109/TPAMI.2004.1273918.
2. Sumner P, Mollon JD. Chromaticity as a signal of ripeness in fruits taken by primates. *J Exp Biol*. 2000;203:1987–2000.
3. Barbur JL, Rodriguez-Carmona M, Hickey J, et al. *Analysis of European Colour Vision Certification Requirements for Air Traffic Control Officers*. London: CAA (UK) Report, CAP 1429, 2016;1–76.
4. Barbur JL, Forsyth PM, Wooding DS. Colour, effective contrast and search performance. In: Schmid R, Zambarbieri D (eds). *Oculomotor Control and Cognitive Processes*. North Holland: Elsevier Science Pub. B.V., 1991,413–30.

5. Walkey HC, Barbur JL, Harlow JA, et al. Effective contrast of colored stimuli in the mesopic range: a metric for perceived contrast based on achromatic luminance contrast. *J Opt Soc Am A Opt Image Sci Vis*. 2005;22:17–28.
6. Barbur JL, Harlow JA, Plant GT. Insights into the different exploits of colour in the visual cortex. *Proc R Soc Lond B*. 1994;258:327–34.
7. Birch J, Barbur JL, Harlow JA. New method based on random luminance masking for measuring isochromatic zones using high resolution colour displays. *Ophthalmic Physiol Opt*. 1992;12:133–36.
8. Kaiser PK, Boynton RM. *Human Colour Vision*. Washington, D.C.: Optical Society of America, 1996.
9. Treisman AM, Gelade G. A feature integration theory of attention. *Cogn Psychol*. 1980;12:97–136.
10. Barbur JL, Forsyth PM. The effective contrast of coloured targets and its relation to visual search.*Visual Search 2*. Taylor & Francis Ltd, 1992,319–28.
11. Pinker S. Visual cognition: an introduction. *Cognition*. 1984;18:1–63.
12. Yamagishi N, Melara RD. Informational primacy of visual dimensions: specialized roles for luminance and chromaticity in figure-ground perception. *Percept Psychophys*. 2001;63:824–46.
13. Cole BL. Colour blindness and driving. *Clin Exp Optom*. 2016;99:484–7. doi:10.1111/cxo.12396.
14. Sachtler WL, Zaidi Q. Visual processing of motion boundaries. *Vision Res*. 1995;35:807–26.
15. Wright WD. *Researches on Normal and Defective Colour Vision*. London: Henry Kimpton, 1946.
16. Nathans J, Piantanida TP, Eddy RL, et al. Molecular genetics of inherited variation in human color vision. *Science*. 1986;232:203–10.
17. Neitz J, Neitz M, Kainz PM. Visual pigment gene structure and the severity of color vision defects. *Science*. 1996;274:801–4.
18. Barbur JL, Rodriguez-Carmona M. Variability in normal and defective colour vision: consequences for occupational environments. In: Best J (ed).*Colour Design*. Cambridge: Woodhead Publishing Limited, 2012;24–82.
19. Squire TJ, Rodriguez-Carmona M, Evans AD, et al. Color vision tests for aviation: comparison of the anomaloscope and three lantern types. *Aviat Space Environ Med*. 2005;76:421–29.
20. Brainard DH, Roorda A, Yamauchi Y, et al. Functional consequences of the relative numbers of L and M cones. *J Opt Soc Am A Opt Image Sci Vis*. 2000;17:607–14.
21. Wyszecki G, Stiles WS. *Color Science – Concepts and Methods, Quantitative Data and Formulae*. John Wiley & Sons, 1982.
22. Barbur JL, Rodriguez-Carmona M, Harlow JA, et al. A study of unusual Rayleigh matches in deutan deficiency. *Vis Neurosci*. 2008;25:507–16. doi:10.1017/S0952523808080619.
23. Zhaoping L, Carroll J. An analytical model of the influence of cone sensitivity and numerosity on the Rayleigh match. *J Opt Soc Am A Opt Image Sci Vis*. 2016;33:A228–37. doi:10.1364/JOSAA.33.00A228.
24. Neitz J, Neitz M, He JC, et al. Trichromatic color vision with only two spectrally distinct photopigments. *Nat Neurosci*. 1999;2:884–88.
25. Belcher SJ, Greenshields KW, Wright WD. Colour vision survey using the Ishihara, Dvorine, Bostrom and Kugelberg, Bostrom, and American-Optical Hardy-Rand-Rittler tests. *Br J Ophthalmol*. 1958;42:355–9.
26. Bergman H, Duijnhouwer F. Recognition of VDU presented colors by color defective observers. *Proc Hum Fact Ergon Soc Annu Meet*. 1980;24:5.
27. Ramaswamy S, Hovis JK. Ability of the D-15 panel tests and HRR pseudoisochromatic plates to predict performance in naming VDT colors. *Vis Neurosci*. 2004;21:455–60.
28. Gegenfurtner KR, Sharpe LT. *Color Vision: From Genes to Perception*. Cambridge; New York: Cambridge University Press, 1999.
29. Sharpe LT, Stockman A, Jagle H, et al. L, M and L-M hybrid cone photopigments in man: deriving lambda max from flicker photometric spectral sensitivities. *Vision Res*. 1999;39:3513–25.
30. Barbur JL, Rodriguez-Carmona M, Evans S, et al. Minimum Colour Vision Requirements for Professional Flight Crew. Recommendations for New Colour Vision Standards. Paper 04/2009, 2009.
31. CIE, ed. Proceedings 1931; 1931. Cambridge University Press.
32. Roorda A, Williams DR. The arrangement of the three cone classes in the living human eye. *Nature*. 1999;397:520–22.
33. Regan BC, Reffin JP, Mollon JD. Luminance noise and the rapid determination of discrimination ellipses in colour deficiency. *Vision Res*. 1994;34:1279–99.
34. Rodriguez-Carmona M, O'Neill-Biba M, Barbur JL. Assessing the severity of color vision loss with implications for aviation and other occupational environments. *Aviat Space Environ Med*. 2012;83:19–29.
35. Barbur JL. 'Double-blindsight' revealed through the processing of color and luminance contrast defined motion signals. *Prog Brain Res*. 2004;144:243–59. doi:10.1016/S0079-6123(03)14417-2.
36. Barbur JL, Connolly DM. Effects of hypoxia on color vision with emphasis on the mesopic range. *Expert Rev Ophthalmol*. 2011;6:409–20.

37. Barbur JL, Rodriguez-Carmona M. Color vision changes in normal aging. In: Elliott AJ, Fairchild MD, Franklin A. *Handbook of Color Psychology*. Cambridge: Cambridge University Press, 2015;180–96.
38. Carroll J, Neitz J, Neitz M. Estimates of L:M cone ratio from ERG flicker photometry and genetics. *J Vis*. 2002; 2:531–42.
39. He JC, Shevell SK. Variation in color matching and discrimination among deuteranomalous trichromats: theoretical implications of small differences in photopigments. *Vision Res*. 1995;35:2579–88.
40. Rauscher FG, Chisholm CM, Edgar DF, et al. Assessment of novel binocular colour, motion and contrast tests in glaucoma. *Cell Tissue Res*. 2013;353: 297–310. doi:10.1007/s00441-013-1675-x.
41. Vemala RV, Sivaprasad S, Barbur JL. Early detection of colour vision loss in age related maculopathy. *Invest Ophthalmol Vis Sci*. 2012;53.
42. Moro SI, Rodriguez-Carmona ML, Frost EC, et al. Recovery of vision and pupil responses in optic neuritis and multiple sclerosis. *Ophthalmic Physiol Opt*. 2007; 27:451–60. doi:10.1111/j.1475-1313.2007.00501.x.
43. Barbur JL, Moro S, Harlow JA, et al. Comparison of pupil responses to luminance and color in severe optic neuritis. *Clin Neurophysiol*. 2004;115:2650–8.
44. O'Neill-Biba M, Sivaprasad S, Rodriguez-Carmona M, et al. Loss of chromatic sensitivity in AMD and diabetes: a comparative study. *Ophthalmic Physiol Opt*. 2010;30: 705–16. doi:10.1111/j.1475-1313.2010.00775.x.
45. Ballard J. Colour-vision safety on the track. *Occup Health Work*. 2013;10:20–3.
46. CIE. Colours of signal lights. Vienna Austria. Standard: S004/E-2001: Bureau Central de la CIE 2001.
47. Cole BL, Maddocks JD, Sharpe K. Visual search and the conspicuity of coloured targets for colour vision normal and colour vision deficient observers. *Clin Exp Optom*. 2004;87:294–304.
48. Birch J. Efficiency of the Ishihara test for identifying red-green colour deficiency. *Ophthalmic Physiol Opt*. 1997;17:403–8.