Colour vision assessment for maritime navigational lookout:

review for UK Maritime and Coastguard Agency (MCA)

Tim Carter, Norwegian Centre for Maritime Medicine, Bergen
John Barbur, Applied Vision Research Centre, City University London

July 2015

Executive Summary

Colour vision in seafarers has to be tested to ensure that all applicants who pass can carry out visually demanding, colour-related lookout duties with the same accuracy as those with normal trichromatic vision, while at the same time there should be no denial of career opportunities to those who can safely carry out such duties. The current test protocol used by MCA, which is in line with the requirements in international conventions, uses the Ishihara test as initial screen, with access to confirmatory testing for those who fail the initial screening.

The Holmes Wright B lantern (HW-B), used for confirmatory testing of colour vision in UK seafarers with navigational lookout duties, is no longer manufactured and servicing cannot be guaranteed. This report was commissioned to evaluate options for its replacement with an alternative test.

The recommended alternative is the CAD (Colour Assessment and Diagnosis) test. This measures the severity and type of colour vision loss, and reliably detects congenital deficiency. When the upper pass threshold limit is set at 2.35 CAD units, the test can provide a valid alternative to the current lantern test. An additional benefit would be a reduction in the number of seafarers with adequate colour vision who are now classified incorrectly as unfit. It is also practical to introduce and to use the CAD test, as it is currently available at a number of centres in the UK.

Recent evidence about the characteristics of the Ishihara test, used as the initial screen for colour vision in seafarers, indicates that the validity of the current testing protocol could be improved by making changes either to the criteria used for failure on this test or by using the CAD test more widely as the primary means of testing. These changes have the potential to ensure that all those with colour deficiencies are identified correctly.

All current approaches are limited by the lack of up to date information on the colour vision requirements for present day navigational lookout duties and the best means of assessing the most colour critical tasks. This means that the relative risks to decision taking while performing such duties in those with normal and with colour vision deficiencies cannot be established

Index

Authors: conflicts of interest

Prof. Tim Carter was formerly Chief Medical Adviser to the UK Maritime and Coastguard Agency, who commissioned this review. He has also acted as adviser to the International Maritime Organization on colour vision testing in seafarers and he co-organised the specialist workshop on the topic held in Kobe, Japan in 2014.

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

Background

1. The MCA requested this review because the current test protocol for colour vision, based on initial screening with Ishihara plates and then confirmatory testing using the Holmes Wright B (HW-B) lantern is becoming unsustainable as the lanterns are no longer manufactured and quality assured servicing arrangements can no longer be reliably obtained. The review aims to answer three questions posed by MCA:

1. "How efficient/appropriate is the current MCA colour assessment protocol?
2. Do CAD or other computerised colour vision tests provide a simple, more accurate and economical alternative to the lantern test currently in use to establish fitness in deck department personnel who fail the Ishihara test for colour vision?
3. If so what cut off threshold limits for the CAD or other tests represent limits that are 'equivalent' to the outcome of the current testing protocol."

2. At the same time it has become apparent, from discussions with maritime authorities and with those who carry out colour vision testing, that there are considerable inconsistencies between the colour vision test methods and protocols used in different countries. The International Maritime Organisation wishes to see this addressed but, as yet, has not been able to mobilise the resources needed to develop a valid and consistent approach that can be adopted internationally. Current testing methods and protocols employed in various countries produce inconsistent outcomes that in some cases are likely to result in unjustified discrimination in employment. In other cases, applicants with congenital colour deficiency may be passed as safe without guaranteeing in any way that they are able to carry out the most demanding safety-critical, maritime colour-related tasks as well as those with normal trichromatic vision.

History

3. The need for colour vision testing in seafarers with navigational duties arose in the nineteenth century, not from the recognition of safety risks in the natural maritime environment but because of the use of red and green lights to indicate the position of ships, and later for markers in seaways. Red and green were the only two feasible colours at a time when oil lamps with wicks, burning with a yellow flame, were used. Co-incidentally, the commonest form of congenital colour deficiency involves red/green colour discrimination. Such deficiencies vary in severity from almost normal colour vision to a total absence of red-green colour discrimination and are found in just over 8 per cent of males and some 0.4 per cent of females.[1]

4. Colour vision testing was introduced, both for seafarers and in the rail industry, in the 1880s after a number of serious incidents linked to the inability to recognise coloured signals or navigation lights correctly. At first these procedures were based on colour naming, but this was soon replaced by colour matching tests. Many anomalous results were obtained, often to the detriment of

the careers of navigating officers. This led to public concern and, in the UK, to a careful study of the validity of different methods of colour vision testing in 1911.[2]

5. A standard colour-testing lantern was introduced in 1914 that was designed specifically to simulate ship navigation lights at a distance. Lantern testing carried out at Board of Trade Marine Offices became part of the assessment of officers thereafter. The current Holmes-Wright B lantern is a linear descendent of the 1914 version. Later developments included the extension of colour vision testing from navigating officers to ratings in the deck department and the use of the Ishihara plates (IH) as a screening test, with only those who failed them being referred for lantern testing.[3] Both the screening test and the confirmatory tests assess colour vision, but in differing ways.

6. The outcomes of both the screening test and the confirmatory tests depend on the applicant's chromatic sensitivity, although the state of light adaptation and the stimulus conditions differ significantly. As a result even those with normal trichromatic vision can on occasions produce inconsistent results on the different tests.[4] This is largely because the visual tasks required of those tested differ (from identifying large characters on a plate when adapted to daylight, to naming correctly the colours of pairs of small signal lights when seen under unfavourable conditions, such long viewing distances, at night). It is well established that visual parameters such as the dimensions of the colour stimulus and the state of light adaptation of the eye affect the ability to see colour differences. This approach is therefore inherently likely to lead to inconsistent findings when the results of different tests are compared.

7. Other maritime nations introduced their own procedures; most were broadly comparable to those in the UK. However the plate tests and protocols used for screening, the tests used on those who had failed at screening and the extent to which there was discretion based on expert, ophthalmological, opinion varied greatly.

8. The CIE (*Commission Internationale de L'eclairage* /International Commission on Illumination) is the main international organisation that sets technical standards for illumination, including those for signal lights.[5] In 2001 CIE produced a technical guide to colour vision testing in transport (CIE 143:2001).[6] The current UK test procedures were and are in line with these recommendations, as were those of a number of other maritime nations, but there were also many national criteria that were at variance with them. However this did not become apparent until the detailed criteria for the issue of medical certificates to seafarers under the International Maritime Organisation (IMO) Standards for Training Certification and Watchkeeping (STCW) Convention became mandatory when it was amended in 2010. This convention indicated that colour vision testing needed to accord with CIE 143:2001.[7]

9. Concerns in some countries about their lack of compliance with CIE 143:2001 only became apparent after the Convention had been agreed and this led to IMO agreeing to delay introduction of the mandatory changes to colour vision testing.

An expert workshop was held in early 2014 that reviewed current practices and the scope for using a number of screen-based tests to replace those listed in CIE 143:2001.[8]

10. The report from the workshop also identified that, as these newer tests could assess the class of colour vision and also measure the severity of colour vision loss, studies were needed to establish appropriate colour thresholds to use to ensure that safety was not compromised without unjustifiably discriminating against those seeking to work at sea. While no recent studies have been done on safety critical colour vision thresholds in the maritime sector, there are no available incident reports suggesting that colour deficiency, even when tested by the least demanding test methods currently in use by some maritime authorities, has contributed to any maritime incidents that have been investigated.

11. However, given that most incidents are multi factorial it is not always easy to assess with confidence the extent to which lack of normal trichromaticity or any other aspect of vision that cannot be classed as completely normal may have contributed to the outcome. Methods for establishing upper threshold limits of colour vision loss that can be classed as safe within specific environments have recently been developed. There is good evidence from other transport sectors, such as aviation and the railways, where tests that can measure the severity and type of deficiency are already in use, that appropriate threshold limits can be derived from studies of safety-critical task demands and that such limits can then form the basis for fair and valid fitness criteria.[9]

Current MCA colour vision test methods

12. The procedures used by MCA have evolved over the years. They were last revised following a review in 1997.[10] Different requirements apply to those in the deck department who may carry out navigational lookout duties (the subject of the current review), those in the engine department who have to recognise colour coded safety information and those who work in catering or other departments of a ship. For crewmembers with navigational lookout duties, initial testing is by Ishihara plates, with anyone who incorrectly identifies more than two plates having the opportunity to then have their colour vision assessed using the HW-B lantern.

13. Approved Doctors (AD) performing statutory medicals on seafarers on behalf of MCA are required to test for colour vision using Ishihara plates at each medical, unless they performed the previous one, in which case test can be done every four years. These requirements are given in Merchant Shipping Notice (MSN) 1839.[11] They are supported by guidance given in Approved Doctors' Guidance (ADG) 14.[12] The relevant sections of both are at Appendix 1. Observation indicates that tests are not performed in exactly the same way by all ADs.

14. HW-B tests are performed at three of the MCA Marine Offices (Southampton, Aberdeen, Beverley). Recently City University London has also been designated

as a centre and has been undertaking comparisons between HW-B, HW-A and CAD results. All testers follow a specified protocol.[13] This requires a blacked out room, which is generally unsuitable for other uses and is only used for testing for a few hours each year. A period of dark adaptation is required prior to testing and preparation for the test and the test itself may take around 30 minutes of a marine surveyor's time.

15. The statistics available from MCA show that in 2014 over 54,000 seafarer medicals were performed. Of these 24 seafarers were failed either temporarily or permanently and 626 were restricted because of colour vision deficiencies. However these crude data do not distinguish between medical examinations at the start of career and those later, nor do they separate out deck, engineering and other crew members: each has different colour vision requirements. As a result they do not enable any useful conclusions to be reached on the importance of the problem, the consistency of test results or their implications for careers at sea.

16. Information has also been collected by MCA on the number of seafarers who have had confirmatory testing using the HW-B lantern because they have failed to meet the MCA Ishihara criteria. Most of these tests will be at the start of career as those who pass the HW-B are given a certificate that means repeat testing is not needed at each medical. Not all those tested are Merchant Navy seafarers as some other groups of maritime workers are also assessed and, in some cases, less demanding criteria for passing are required. The pass rates vary between the four testing centres, either because of differences in the application of the test protocol or because some centres see a higher proportion of those who only need to meet the less demanding criteria.

17. An average of 98 tests have been done each year (2003-2014) with an average of 20 passes. However the City University centre, which has been following the testing protocol exactly as specified for HW-B, had no passes in the tests done since 2012, suggesting perhaps that variable interpretations about the testing procedures have enabled some candidates to pass. Additional tests of colour vision were also performed at this centre (Table 1).

Sample statistics - MCGA applicants that took HW-A, HW-B , lanterns and other tests at City University						
Sample size (21)	N	% pass IH (less than 3 errors)	% pass HW-A	% pass HW-B	% Nagel (MR<9)	% pass CAD (RG)
Normal trichromats	2	100	100	100	100	100
Deutan	14	14	36	0	64	0
Protan	5	20	0	0	100	0
Acquired	0	NA	NA	NA	NA	0

Table 1. Test results from City University on MCA subjects who failed initial Ishihara test. (The terms deutan and protan are explained in paragraph 19 and Figure 1). 24% of those who failed the first Ishihara test (plates 1 to 25) pass on retest. 36% of deutans passed on the HW-A lantern, but none on the HW-B lantern. The CAD test results were fully consistent with the HW-B findings. (CAD and Nagel tests are described in paragraphs 38 and 42-44).

Relevant Vision Science

18. This section reviews those aspects of vision science that are relevant to the assessment of colour vision in seafarers. Colour vision deficiency may be either inherited or acquired.

Mechanisms of colour perception

19. The normal retina contains three classes of cone photoreceptor with distinct spectral responses: short wavelength (S-cones), middle wavelength (M-cones) and the long wavelength (L-cones) (Figure 1). In this section the results from several hundred subjects tested using the CAD test are used as a benchmark. This test is described later in this section.

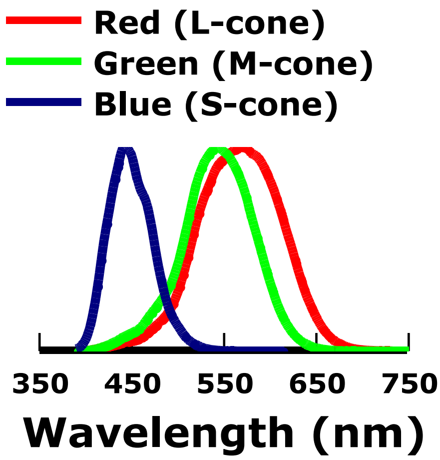


Figure 1. Spectral sensitivities of L (red), M (green) and S (blue) cones in a normal trichromat.

Red / green dichromats lack either the L-cones (protanopes, protans) or the M-cones (deuteranopes, deutans).

Deuteranomalous subjects rely on a variant L-cone pigment that differs in the wavelength of peak spectral responsivity from the normal L-cone pigment by as much as 12nm. [14]

Protanomalous subjects rely on a variant-cone M pigment with a maximum peak separation of ~ 7nm from the normal M-cone pigment.

Subjects with absent or non-functioning S-cones are very rare (~ one in 15 to 20 thousand).[1]

Subjects with acquired loss of colour vision (that can affect both the RG and YB colour vision) are more common above 55 years of age and increase rapidly above 65 years of age (often exceeding the percentage of those with congenital deficiencies). [15]

Inheritance of colour vision deficiency

20. Normal colour vision depends on the presence of each of the three colour sensitive pigments in a form that has the optimum spectral response. An understanding of the pattern of inheritance for colour vision and its expression in terms of limitations to colour discrimination is needed in order to understand the rationale for colour vision testing and to evaluate its ability to assess the nature, severity and significance of any deficiencies identified.

21. The genes responsible for the red / green colour sensitive pigments form a sequential array and are localised to the X-chromosome. A person may inherit a variant gene in the array that determines the spectral tuning (wavelength of

maximum sensitivity) of red or green cone pigments. This will result in anomalous colour vision. There are several cone pigment genes in the array (e.g., L, M, M, L or L, L, M). Although the first two genes are most likely to be expressed sometimes even the third one can be expressed) When the first two cone pigment genes in the array are of the same kind, it is highly likely that these are expressed resulting in the same pigment class which results in dichromacy (i.e., the absence of red-green colour discrimination).

22. The proximity of red and green cone genes in the array facilitates genetic mutations and hence the production of variant pigments. The small wavelength separation in peak spectral responsivity of red (L) and green (M) cones (28-30 nm) means that even moderate shifts in peak wavelength sensitivity may cause significant changes in red-green colour vision.

23. The red and green cone pigment genes are carried on the X chromosome hence their expression can be expected to differ in males and females. In females there are two of these chromosomes and, provided one is fully functional, red and green pigments will be produced. If the cone pigment genes on the first chromosome produce only red cone pigment and the second chromosome only green pigment, functionally normal colour vision will result although the relative numbers of red and green cones may differ significantly from the 2:1 ratio observed in those with normal trichromatic vision.

24. Males only have one X chromosome and so if both cone pigment genes in the array specify the same pigment type there will be complete loss of red-green colour vision. If one of the first two cone pigment genes specifies a variant pigment of the same class, the subject will have reduced (anomalous) chromatic sensitivity. It is for this reason that deficiencies are much more frequent in males.

25. The separation between the spectral sensitivity peak of blue (S) cones and green (M) and red (L) cones is large (up to 130nm) and so anomalous variants of blue pigments are unimportant. Deficiency of this pigment is very rare. The majority of deficiencies in blue perception are therefore linked to acquired colour deficiency (i.e. a loss of chromatic sensitivity caused by age-related eye disease).

Acquired loss of chromatic sensitivity

26. The ability to discriminate or see small colour differences can be affected severely by diseases of the retina and the visual pathways. Eye disease is associated with loss of both yellow/blue and red/green discrimination and in older people this can be an even more important contributor to colour vision deficiency than congenital deficits. Figure 2 shows the contribution of various conditions to the increased colour thresholds. The subjects were all refracted appropriately. Small amounts of refractive error do not affect the CAD test results. Traditional yellow/blue colour vision has been considered less relevant in safety critical jobs. The relatively recent introduction of blue LED-based signal lights and the extensive use of yellow and blue colours on modern visual displays

means that yellow/blue colour vision loss, usually from acquired deficiency, has now become an important part of colour vision assessment for some safety critical tasks.

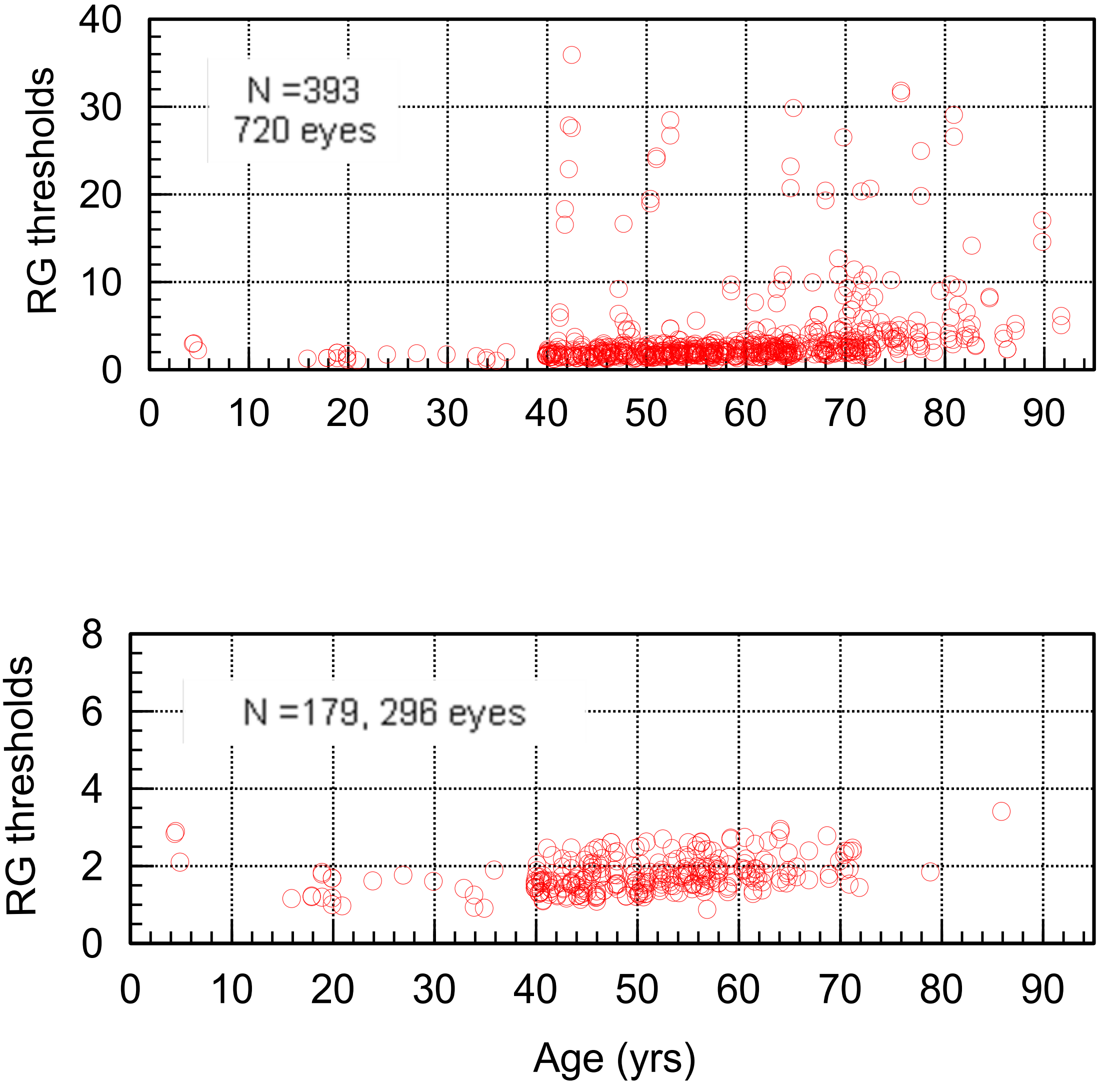


Figure 2. The upper graph shows CAD thresholds measured in 393 subjects who attended an optometric clinic for routine refraction and eye checks (See paragraph 42 for explanation of CAD thresholds and units). Acquired red/green deficiency is important and not uncommon in older subjects. The lower graph shows the remaining 'normal' eyes after excluding the following conditions:

- Subjects with congenital colour deficiency (i.e. elevated RG thresholds and 'normal' YB)
- Subjects with medical conditions such as diabetes, hypertension and ocular abnormalities
- Subjects with abnormal fundus appearance or drusen
- Subjects with statistically significant differences (in either RG or YB or both) between the two eyes.[16]

Effects of normal aging on red-green and yellow-blue colour vision

27. Colour perception sensitivity declines with age (Figure 3). This occurs even in the absence of other eye disease.[15] The effects of normal aging on RG and YB colour sensitivities are relevant to the frequency of testing and to which forms of test are most effective at identifying those with safety critical deficiencies, especially later in their careers when eye disease is likely to be more frequent.

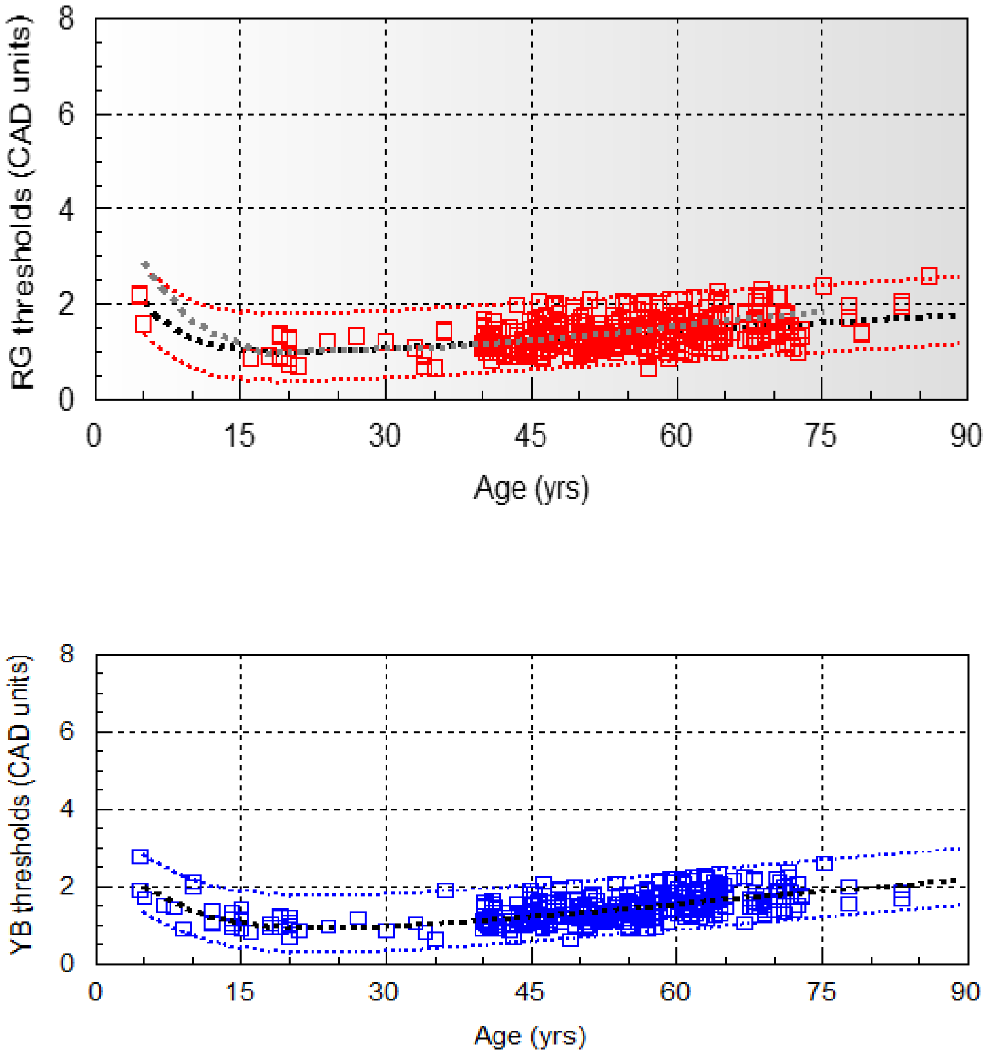


Figure 3: RG (A) and YB (B) binocular thresholds as a function of age for the normal trichromats examined in the study and deemed to reflect the effects of normal aging on colour vision (See paragraph 42 for explanation of CAD thresholds and units) The filtered results show significant, but relatively uniform variability and only a gradual increase in colour thresholds as a function of age. black dashed lines in sections A and B describe how normal aging affects colour vision and the corresponding dotted lines above and below the mean represent the $\pm 2.5\sigma$ limits. [15] [Koenker's test for heteroscedasticity carried out for data points within age range 35 to 75 years confirmed this observation (H_o-true, p=0.224, n=270).[17]
The mean binocular thresholds as a function of age are given by:
RG_{bin} = 0.698 + 0.0121*age +3.373*exp (-0.19*age) and YB_{bin} = 0.24 + 0.0218*age +2.99*exp (-0.1136*age).]

Characteristics of colour vision test methods

28. Most currently used colour vision tests aim to detect red/green deficiencies because of their high frequency and functional importance. Some tests produce a measure of the severity of colour vision loss but the majority use arbitrary cut off points as the basis for determining the presence or absence of colour vision deficiency.

Ishihara plates

29. The most widely used test is the Ishihara book of plates (IH). The plates improve the identification of colour vision deficiencies by using several different approaches such as confusion between and disappearance of digits as well as digit recognition and trail following. The number of plates the subject fails to read correctly does not, however, provide a reliable indication of the severity of red/green colour vision loss.[18] When applied correctly some 19% of normal trichromats and almost all colour deficient observers (98% of deutan (green deficient) and 99% of protan (red deficient) subjects) fail one or more plates on the Ishihara test (Figure 4).[19] These findings are confirmed from a number of studies on occupational populations (Table 2).

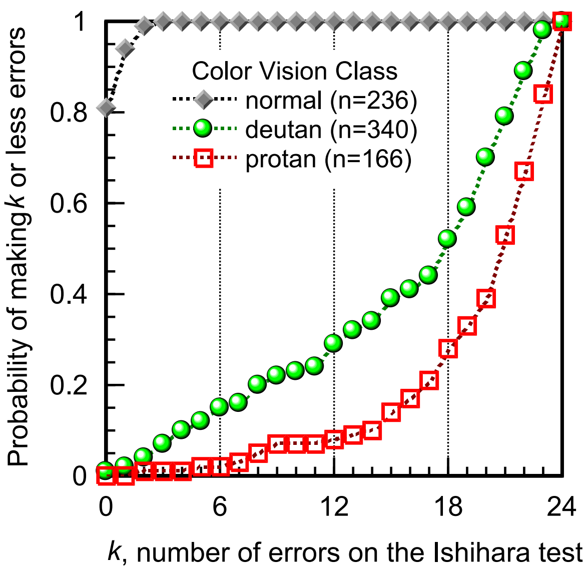


Figure 4. The probability of making (*k*) or less errors when reading the numerals on the Ishihara test plates plotted for a group of normal trichromats and for subjects with congenital, deutan and protan colour deficiency. The order of presentation was random for the remaining 24 plates of the Ishihara 38-plate test. [16]

Table 2	CAA		FAA			TfL	Fire service
Edition	38-plate ed	24-plate ed	38-plate ed	24-plate ed	14-concise	38-plate ed	24-plate ed
Pass Criteria	1-25 plates no errors	1-15 plates no errors	1-21 plates <=8 errors	1-15 plates <=6 errors	1-10 plates <=5 errors	1-17 plates no errors and ≤ 3 esp*	1-17 plates <=2 errors
N (236)	191	213	236	236	236	229	235
D (340)	2	5	68	64	108	5	24
P (166)	0	0	10	10	14	1	2
% N	80.93	90.25	100.00	100.00	100.00	97.03	99.58
% D	0.59	1.47	20.00	18.82	31.76	1.47	7.06
% P	0.00	0.00	6.02	6.02	8.43	0.60	1.20

*esp = errors made on specified plates

Table 2. Data showing the percentage of subjects within each colour vision class that pass the Ishihara test for the different, currently accepted testing protocols within the listed professional environments. The subject groups represent random samples that reflect normal (N), deutan (D) and protan (P) colour vision classes. The samples do not represent the distribution of colour deficiency within any occupational population.
Before the introduction of the CAD test the CAA (UK Civil Aviation Authority) used both the 38 and the 24 plate editions to screen for pilots and air traffic controllers. JAR guidelines (for European commercial pilots) allow for the use of the first 24 plates of the 38 plate edition. The USA Federal Aviation Administration (FAA) also use the 38 plate edition, but only use 21 plates. In addition, the FAA also allow a number of other different tests. Currently, TfL (Transport for London) use only the CAD test. Fire service applicants are tested at City University using the CAD test. [20]

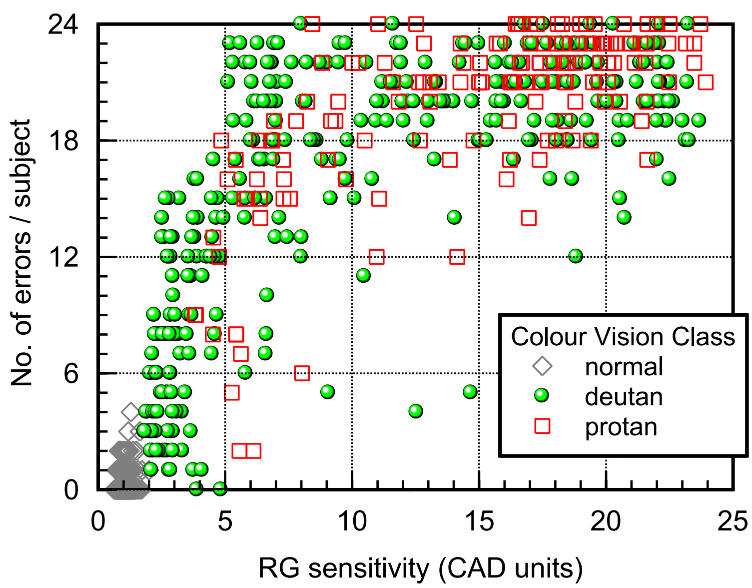


Figure 5. The number of errors made by each subject on the Ishihara test plotted against the subject's RG threshold (measured in Standard Normal (SN) CAD units. Data are shown for normal trichromats (grey diamonds), deutan (green discs) and protan (red squares) subjects. [19]

30. The frequency of errors on Ishihara plates can be compared with the grading of severity of colour vision loss using more accurate tests for colour thresholds (such as the CAD test, paragraphs 42-44). This shows only a weak correlation

between the number of plate errors made and the severity of the deficiency (Figure 5).

It can be seen that if no errors on the Ishihara test are allowed almost all those with congenital colour deficiency fail, but 19% of normal trichromats also fail. By contrast if 2 or less errors are accepted more normal trichromats pass although in order to pass all normals one needs to allow 4 or less errors. This error score also passes 10% of deutan subjects and 1% of protans.

Classification of class of colour vision using Ishihara				
	Normals (n=340)	Deutans (n = 343)	Protans (n=169)	Tritans (n=3)
% classified as normal	81	1	0	100
% classified with CCD	19	99	100	0
% classified as deutan	0	62	6	0
% classified as protan	0	0	48	0
% classified as indeterminate	19	38	46	0
% classified as tritan	0	0	0	0

Table 3. Ishihara test (1 to 25 plates), all plates read correctly, in 855 subjects (Normals (340), Deutans (343), Protans (169 and Tritans (3). Note that this plate test does not test for yellow-blue loss, hence the tritan subjects are unclassified.
The results show that the test yields poor classification with only 62% and 48% classified correctly as deutan and protan, respectively. Interestingly, 19% of normals are classed as defective and unclassified whilst 6% of protans are classified as deutan. (CCD – congenital colour deficienciy)

31. The Ishihara isochromatic plate test, when used as a screening test, has the advantage of ease of use and low cost, but the test lacks specificity if sensitivity is to remain high and also fails to diagnose accurately the class of deficiency involved. Table 3 shows the classification for normal, deutan, protan and tritan subjects based on the Ishihara test. Use of this test as the sole criterion for suitability for navigational lookout duties, if no incorrect readings are accepted, means that a number of those with normal colour vision are denied employment, whereas if some errors are accepted, as seen in Figure 5, not all those with colour vision loss will be correctly identified.

32. Because of these characteristics neither Ishihara nor other pseudo-isochromatic plate based tests can be optimised for use as a single definitive test for fitness to undertake navigational lookout duties, which will ensure suitability without denying some of those tested employment. It is for this reason that a second ‘confirmatory’ test is needed for those who do not achieve the criterion set for the Ishihara. The criterion currently set for the Ishihara by MCA will reliably let almost all of those with normal colour vision pass, but will fail to detect some of those with deficiencies in red/green colour vision (Figure 5 shows the frequency and severity of reductions in colour threshold that will arise from acceptance of different numbers of Ishihara plate errors). This latter group will avoid further scrutiny, but like others at the borderline of normal vision they are more likely to make errors when repeat Ishihara tests are done as part of their statutory two yearly medicals. The significance of such deficiencies as predictors of errors during lookout duties is not established. The inability of

the Ishihara test to be used to determine the severity or nature of deficiencies has only become apparent as a result of recent studies undertaken since the current MCA test protocols were reviewed in 1998.

Figure 6. Photograph showing vertical section through Ishihara test plate 2 (which displays number 8). Learning experiments carried out with black and white photocopies of the IH test plates show that within two hours a subject can learn to make use of the spatial distribution and size of dots along either the vertical or horizontal sections of the circumference to pass with less than two errors.

33. The Ishihara and other plate tests require appropriate illumination and correct presentation and recording of results plate by plate, as well as a method of use that avoids conscious or unconscious prompting. As the pattern of dots at the margin on some plates can provide candidates with non-colour related clues that can be learnt, examiners need to closely observe the behaviour of those tested (Figure 6). It is possible that testing in poorly standardised clinical settings may not give the same level of reliability as that obtained during carefully controlled research studies.

<u>The Holmes-Wright Lanterns </u>(see Appendix 2)

34. The current confirmatory test, the HW-B lantern, is used only on a group of people pre-selected because they have not met the criteria set for the Ishihara test. The HW-B lantern design was based on assumptions about the luminous intensity of ship navigation lights at a distance of two miles, thus it is to an extent a 'trade test' rather than one that is derived from studies of visual performance. It is, however, a demanding test that, when used according to the current MCA protocol, fails over 10% of normal trichromats and all subjects with congenital colour deficiency. This is discussed in paragraph 47 where the relevance of this to multi stage test protocols is reviewed. It follows that an applicant with mild congenital loss of colour vision or a subject with normal trichromatic vision but with poorer RG colour sensitivity is more likely to pass the Ishihara test plates than the HW-B confirmatory test. The variability in repeat HW-B tests is also high. The more widely used Holmes-Wright A lantern (HW-A) has a larger

aperture than the HW-B but similar colour characteristics: all normal trichromats pass, 1% of protans and 22% of deutans also pass (see Figure 7).

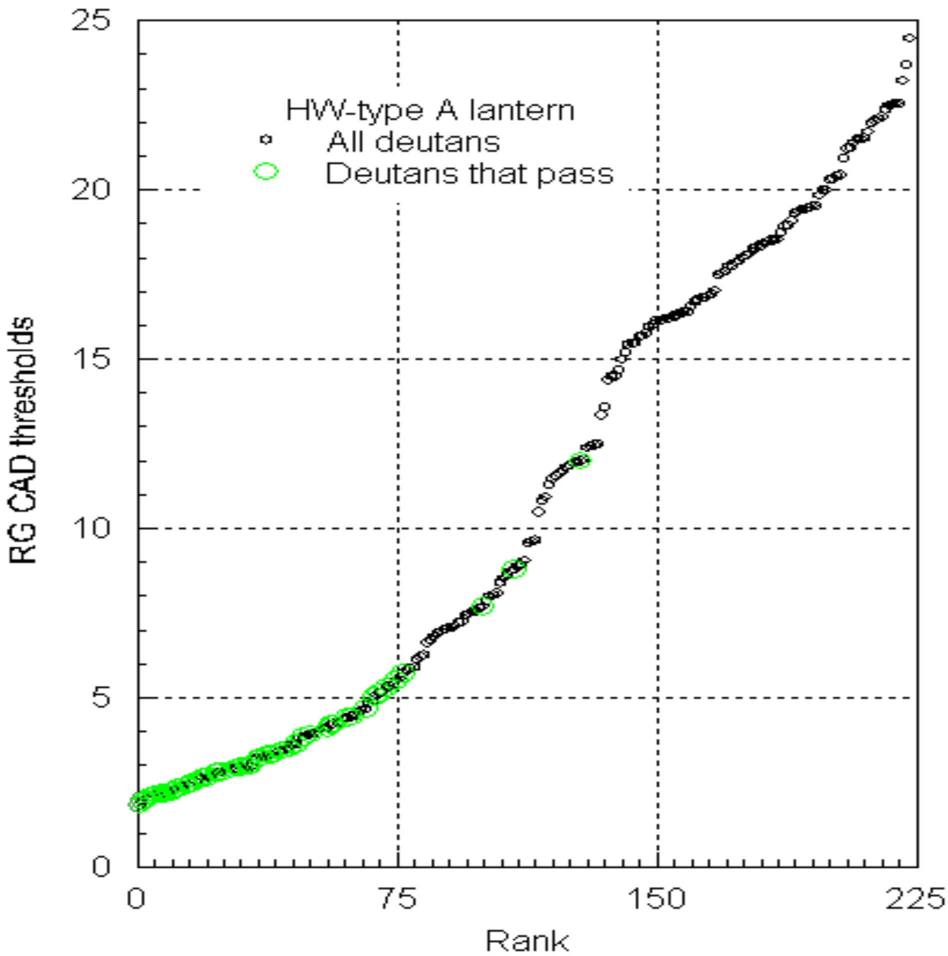


Figure 7. Ranked RG thresholds (CAD units) for 226 deutan subjects investigated. Each subject was also tested on the HW-A lantern using the CAA protocol.

35. Both HW lanterns, at a viewing distance of 6m with aperture diameters of 0.87'(type A) and 0.3' (type B), tend to produce close to 'diffraction limited' images. The total light flux captured by the eye and hence the perceived brightness of a diffraction-limited image is proportional to the reciprocal of the square of the distance between the signal lights and the eye. The larger size aperture in type A causes an almost 10-fold increase in light intensity and as a result the applicant finds it easier to name the correct colours. This is the equivalent of a reduction in the distance at which the same intensity signal light would be seen from 2 to 0.63 miles. The correlation between the HW-A and HW-B lanterns is discussed at paragraph 50 where this is discussed in relation to sequential test protocols.

36. The HW-A lantern is used in other modes of transport such as aviation there is much more detailed information on its performance than on that of HW-B. This is because users in the aviation and rail industries have funded extensive

investigations of its performance, but similar studies have not been done for the HW-B, which is used only in the maritime sector.

Other lanterns

37. A number of other lanterns are or have been in use such as the Farnsworth lantern (Fallant), the Fletcher CAM, the ALT (FAA Aviation Lights Test: Figure 8) and the Beyne. The ALT lantern is a modified version of the Farnsworth lantern using filters that bring it more in line with the HW-A lantern. The pass / fail rates on the ALT lantern are also very similar to the HW-A lantern. The Fletcher CAM is designed to mimic the HW-A and –B lanterns, but has proved not to be satisfactory in use. Lanterns can differ in terms of their intensity, aperture size and whether their filters provide the same spectral output, as opposed to similar chromaticities. For instance, the Falant lantern employs a 2.54mm aperture viewed from 8 feet (~ 2.44m). The angle subtended by image~ 3.58'; this means that the aperture is resolved by the great majority of subjects and images are not diffraction limited but are seen to have a finite size. These parameters and the protocol employed significantly affect the outcome in terms of passes / failures. A systematic review of the published data on the performance of lanterns has recently been completed.[21] Several lanterns are currently are out of regular production and have problems with quality assured servicing. Therefore further comparisons are not needed.

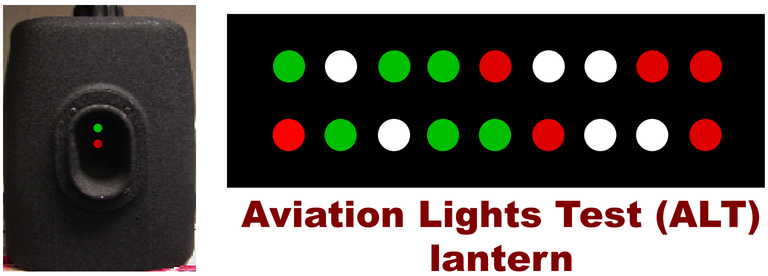


Figure 8. Photograph showing the ALT lantern with vertical pairs of colours very similar to those produced by the HW-A lantern.

Other established test methods

38. Some countries use other test methods as confirmatory tests such as the Nagel anomaloscope test (Figure 9). This instrument can provide valid data on the relative sensitivity of the eyes to red and green light, but neither the match midpoint nor the size of the red/green matching range (MR) provide reliable information on the severity of colour vision loss.[22] This test is an accepted alternative to the HW-B lantern in CIE 143:2001. The anomaloscope is however a costly optical instrument and requires maintenance and a skilled operator in order to obtain reliable and reproducible results.

Figure 9. The Nagel anomaloscope presents a split field stimulus (usually 2 degrees in diameter). This shows a monochromatic yellow light and a red / green mixture field. The subject's task is to adjust the red/green mixture to match the monochromatic yellow field. A normal trichromat accepts only a small range of red / green mixtures as matches to the yellow. In those with colour deficiencies the match points differ and the matching range may be wider. However some of those with unusual genetic mutations can produce matching ranges and match midpoints that fall within the normal range.[23] In addition differences in the optical density (i.e., the amount of visual pigment) between L and M cones can also produce significant shifts in match midpoint, even when the subjects have excellent red / green chromatic sensitivity, but may not be able to make completely normal matches

39. CIE 143:2001 makes provision for the use of a second plate test as an alternative confirmatory test when IH is failed. Most other plate tests are not as well validated as the IH, although some such as the Hardy, Rand and Rittler (HRR), which overall has a lower sensitivity than IH, do have advantages such as a format that makes memorisation of the plates more difficult and an ability to identify yellow/blue as well as red/green deficiencies.

40. The Farnsworth D 15 test is commonly used for those with less demanding colour vision tasks than navigational lookout, but some countries use it for selection for these duties. It is less sensitive and will pass a proportion of those with a significant loss of colour discrimination (Table 4)

D15 analysis			
	P	**F**	**% of subjects that pass**
N(81)	81	0	100%
D(232)	136	96	59%
P(93)	48	45	52%

Table 4: Pass and failure rates on Farnsworth D 15 test for those with normal colour vision (N), deutans (D) and protans (P), showing its low sensitivity for detection of colour vision deficiency.

Screen based tests

41. Several screen-based tests have been developed and marketed for occupational colour vision assessment in recent years.

The CAD Test

42. The best-validated and most widely available screen-based test in the UK is the Colour Assessment and Diagnosis (CAD) test.[24] The test is based on findings from camouflage studies and makes use of dynamic luminance contrast noise to isolate fully the use of RG and YB colour signals (Figure 10). The CAD test has been shown to provide identifiable thresholds that closely correspond to those in current use for HW-A lanterns in aviation (Figure 7).[25] The CAD test has the ability to diagnose accurately the subject's class of colour deficiency and to quantify the severity of colour vision loss. It also separates those with normal trichromatic vision from those with any form of deficiency (Figure 11).

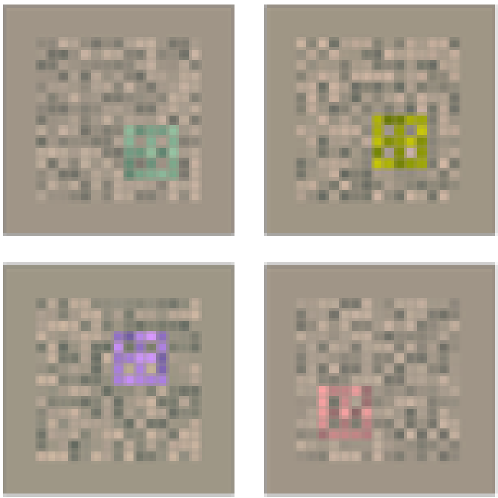


Figure 10: The CAD test screen. A flickering image is presented, through which move blocks of varying colours. The person being tested has to identify the direction of travel of the coloured blocks. The colour intensities are then adjusted depending on the subject's response accuracy and this allows the thresholds for detection of each colour to be determined.

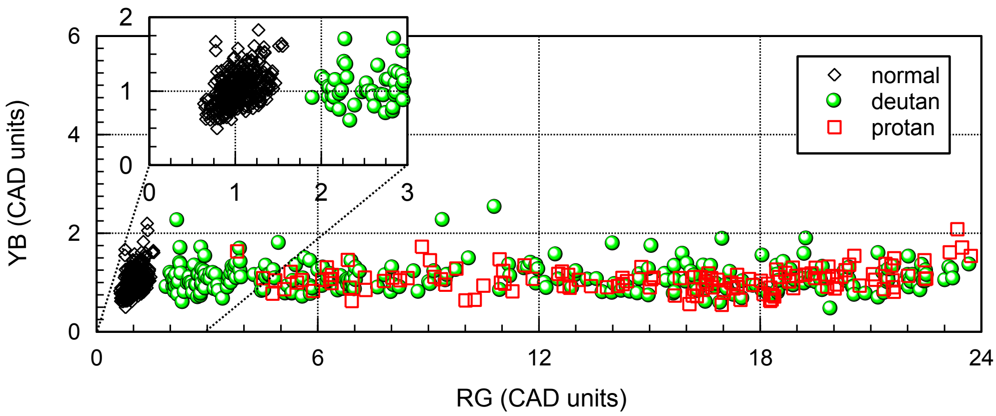


Figure 11. CAD red / green and yellow / blue thresholds for normal trichromats (black symbols), deutan (green symbols) and protan (red symbols). The inset shows a magnified region of the full graph to show the clear separation between normal trichromats and subjects with acquired colour deficiency.

43. Performance is measured in CAD Units (Figure 12). These units are derived from population studies and one unit represents the average signal level needed for colour perception in a young observer (15-35) as their colour discrimination is tested in both red/green and yellow/blue axes. This approach makes it easy to

quantify and understand the severity of colour vision loss. The CAD test uses commercial software and should be used with a high quality monitor with full spectral and luminance calibration. The output display provides a clear summary of the results and their implications for the task being tested for (Figure 13).

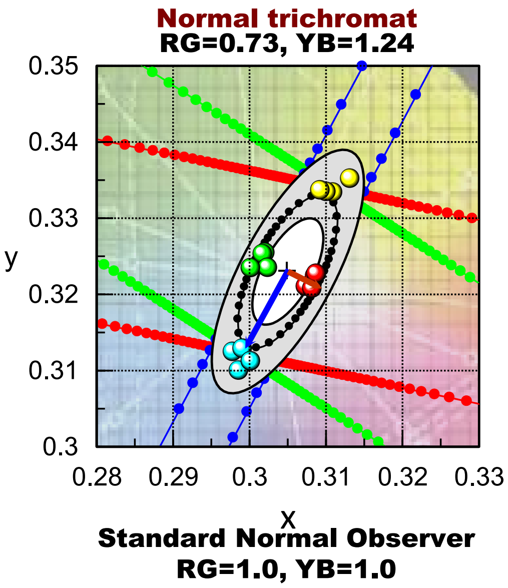


Figure 12. Example of CAD test results for a young, normal trichromat. The results are expressed in CAD units and are plotted on the standard CIE Chromaticity chart. One CAD unit describes the colour signal strength at threshold for young normal trichromats obtained by averaging the thresholds measured in 333 subjects. The data obtained in this way also provide statistical limits that define variability within normal trichromatic vision. These limits are indicated by the grey shaded area bounded by the ellipses. The centre, black-dotted ellipse shown the median threshold which define the RG and YB CAD limits. A subject with a threshold of 3 CAD units requires a three times larger colour signal strength when compared to the median, normal young observer. Data showing the effect of normal aging on RG and YB thresholds (as measured with the CAD test) are shown in Fig. 3

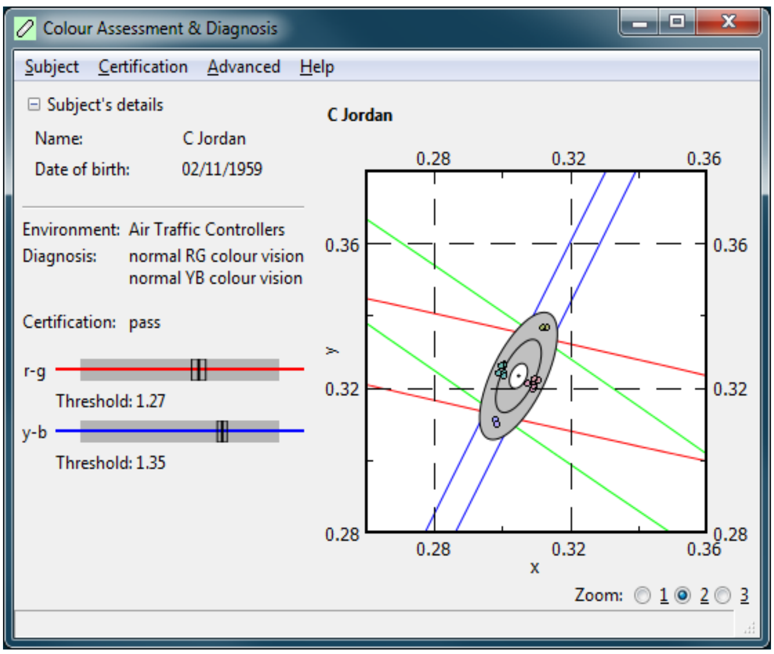


Figure 13. CAD output window showing the applicant's RG and YB threshold. (C. Jordan is a pseudonym)The horizontal, grey bars on the left (under Certification) show the limits of the normal range for the applicant's age, together with the measured thresholds. The centre dots show the actual thresholds measured along the RG and YB colour directions. The graph provides a full summary of the results of the CAD test. Instructions on testing procedures are included with the majority of Menu functions. .Pass / fail outcome based on task specific criteria are also provided for some working environments.

44. The calibration of the display needs periodic checking, usually every six months, and the photometer then recalibrates the display automatically. The test does not require dark adaptation. The test starts with a learning period, usually taking less than one minute. For occupational use this is followed with the Rapid Screening option which tests for both RG and YB colour vision in just over one minute. When the age-corrected version of the CAD test is used ~ 95% of normal trichromats pass and all congenital colour deficient fail (~ 8%). Hence ~ 12.6 % of applicants will require the full RG CAD test, taking around ten minutes, when first investigated. The YB test takes around 3 minutes to complete. The results of the full CAD test (RG and YB) define the applicant's class of colour vision (i.e., normal trichromacy, deutan-, protan- or tritan-like deficiency of congenital origin). In addition the test also identifies acquired loss of colour vision and quantifies the severity of colour vision loss.

<u>The Cambridge Colour Test (CCT)</u>

45. The CCT employs static noise to isolate the use of colour signals and in this respect is similar to the Ishihara and / or the American Optical HRR test, except for the use of a black and not a white background (Figure 14). The test measures unidirectional colour thresholds along the three colour confusion axes or chromatic detection ellipses (Figure 15). This test has mainly been used in research studies and has not been validated in large scale studies for occupational fitness assessment.

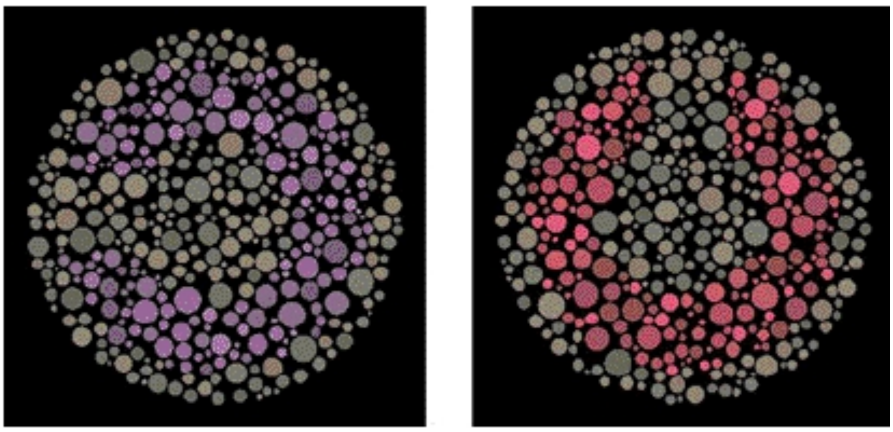


Figure 14: Examples of CCT test images displaying a Landolt ring made up of coloured discs of varying luminance. The subject's task is to indicate the orientation of the gap. [26]

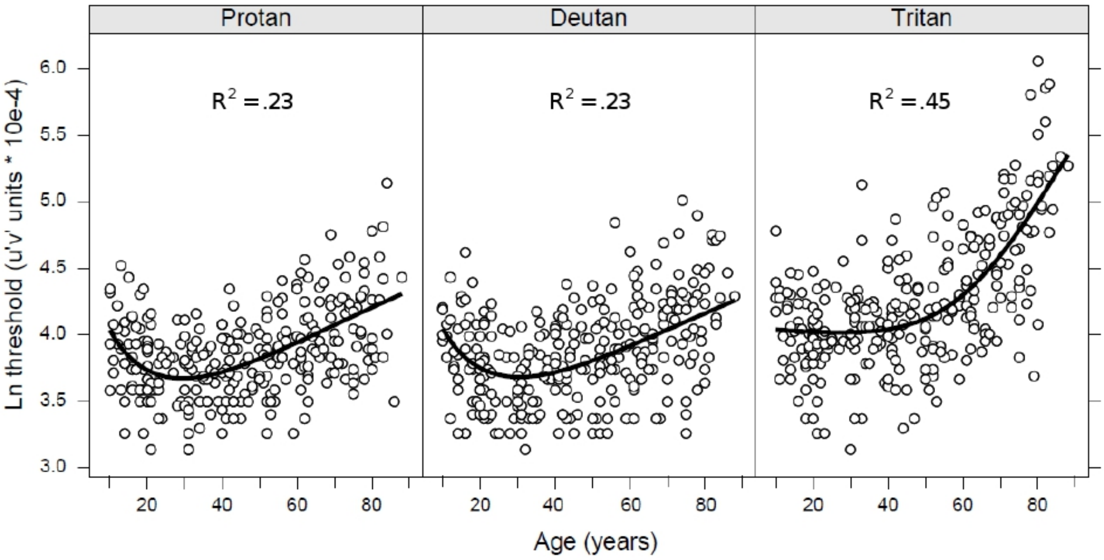


Figure 15: Unidirectional thresholds measured using the Cambridge Colour Test along colour directions that correspond to the colour confusion axes of deutan, protan and tritan observers. The thresholds were measured in subjects with normal trichromatic vision and are plotted as a function of age. The threshold is measured as a displacement away from background chromaticity in the CIE – (u' v') chromaticity chart.

Cone Contrast Sensitivity Test and Conan DX

46. Two other tests, the Cone Contrast Sensitivity Test (CCST) and the Conan DX have been assessed in smaller groups of USA military flight crew but not more widely.[27] [28] The cone contrast test aims to isolate and assess L, M and S cones separately, however as discussed in paragraphs 20-25 this is a complex task in subjects with congenital colour deficiency. The Conan DX test is similar to the CCT and Ishihara in that coloured numbers are produced as dots that vary randomly in luminance so as to minimise the use of luminance contrast in reading the numbers, particularly when colour deficient subjects are involved. Both tests have the advantage of lower initial purchase cost compared with the CAD and CCT tests, but the limited validation data on the performance of these two tests as well as the CCT makes them unsuitable for adoption at present. A programme involving a large amount of assessment in a wider range of populations as well as comparisons with other test methods is needed before they can be considered. This would take some years and considerable resources.

Testing protocols and use of sequential tests

47. Test protocols should be based on information about the characteristics of the available test methods. Each test can be reviewed on its own (Table 5), but a more useful analysis for the purposes of this review is to compare the performance of those who pass and those who fail the MCA Ishihara testing protocol on other tests (Figures 16-18, paragraphs 49, 50).

Sample statistics - subjects that took HW-A and HW-B lanterns and other tests						
Sample size (116) Males (104) Females (12)	N	% pass IH (<3 errors)	% pass HW-A	% pass HW-B	% Nagel (MR<9)	% pass CAD (RG)
Normal trichromats	38	100	100	87	100	100
Deutan	49	6	33	0	55	0
Protan	26	0	4	0	12	0
Acquired	3	100	100	67	100	0

Table 5. Comparison of outcome using Ishihara, HW-A, HW-B, Nagel and CAD tests. 116 subjects were investigated using the Ishihara tests (using plates 1 to 25).

48. However, no test will be a fully reproducible substitute for any other test as it will be performed in different conditions of lighting and will also frequently assess different uses of colour signals. For instance a screen-based test, such as the CAD, will be free from confounding factors such as test stimulus reflectance, changes in the spectral power distribution of the illuminants, aperture size, acuity limits, state of dark adaptation or ability to name colours. In spite of these advantages, screen-based tests will not, however, reproduce some of the subsidiary features of a 'trade test' such as HW-B. These differences in outcome will also apply to related test methods. As a result, even lantern tests that have been designed to mimic the properties of the HW lanterns (e.g., ALT and CAM lanterns) do not give identical results.

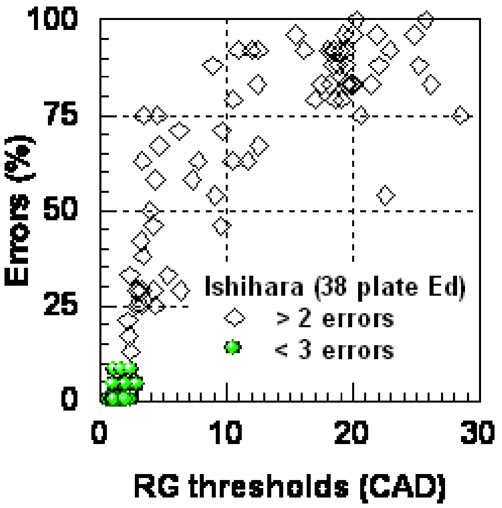
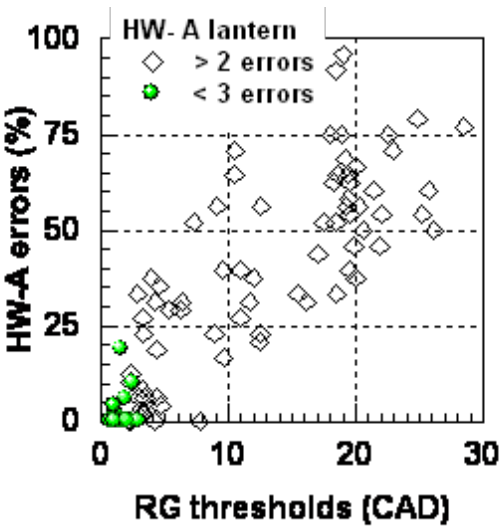


Figure 16: Percentage Ishihara test error scores (using plates 1 to 25) plotted against the subject's RG CAD 117 subjects were investigated (Normal = 45; Deutan = 46 ; Protan = 26). The green symbols show subjects with less than 3 errors. Subjects with 3 or more errors are shown in black. 72 subjects failed with 3 or more errors Subsequent figures use this population to explore the comparability of different confirmatory tests.



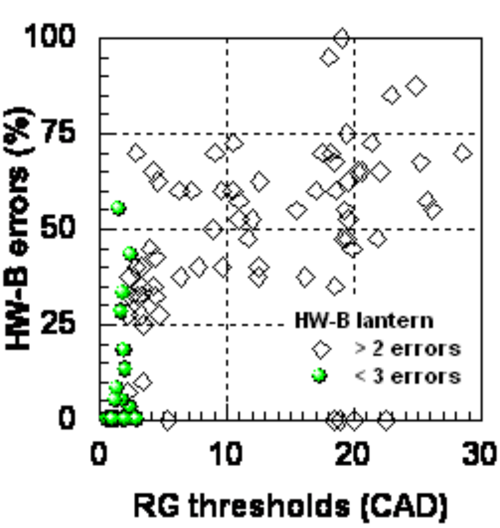


Figure 17: Performance of those who passed (green circles) and those who failed MCA Ishihara protocol on HW-A and HW-B lanterns, with reference to CAD thresholds. It can be seen that, while HW-A errors are rare in those who pass the Ishihara there is a significant failure rate on HW-B for those who pass Ishihara and have minimal loss of CAD threshold.

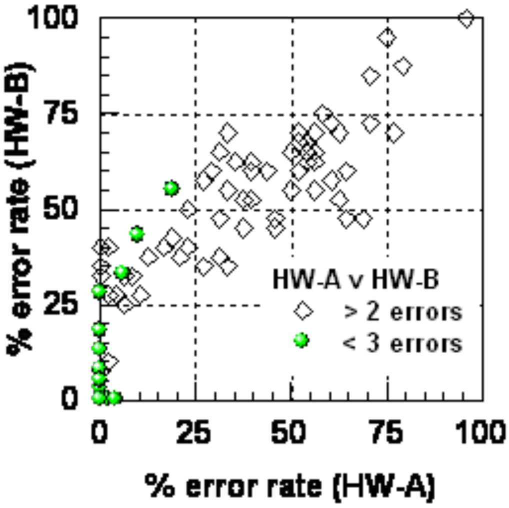


Figure 18: Comparison of HW-A and HW-B results in those who passed and those who failed MCA Ishihara protocol. This indicates that, while the error rate on HW-A is compatible with that from Ishihara, the HW-B is a considerably more demanding test and has a higher rate of misclassification. If the misclassification of those who have normal colour vision and the mild deutans who pass the MCA Ishihara protocol by the HW-B is discarded then the results of the two tests are comparable.

49. The implications of sequential testing and its results using different confirmatory tests are shown in Figures 16-17. The CAD RG threshold is used for reference in Figures 15 and 16. Figure 18 shows the correlation between HW-A and HW-B lanterns. The results are of interest since they also show that normal and mild deutan subjects that pass HW-A can end up with HW-B error scores in the range 0 to 30%.

50. The correlation between the HW-A and HW-B error scores is consistent for applicants who pass using the MCA Ishihara protocol (HW-B results can be predicted well from subject's HW-A results: HW-B (% error score) = 27.502778 + 0.6491888*HW-A ($r^2 = 0.72$). The intercept shows that on average the subject makes 27% more errors on the HW-B lantern. The latter is a demanding test which according to the current MCA protocol fails just over ~ 10 % of normal trichromats and all subjects with congenital colour deficiency. It follows that an applicant with mild congenital loss of colour vision or a subject with normal trichromatic vision but with poorer red/green colour sensitivity is more likely to pass the IH test plates than the HW-B confirmatory test. The variability in repeat lantern tests is also high. The consistency of the gradient between HW-A and HW-B results, once the intercept caused by the over-sensitivity and misclassification when using HW-B is discarded, means that correlation data from the extensive cross-validation studies between HW-A and CAD can be used to provide valid indicators of the degree of comparability between HW-B and CAD results.

51. Any two stage process will only be effective at correctly identifying those with colour vision deficiencies if the threshold for the initial screening test is set at such a level that all such deficiencies are identified. For Ishihara this means using zero plates incorrectly read. However a consequence of this, as distinct from the current MCA practice of accepting a small number of incorrect readings,

will be that a larger number of those with normal colour vision are referred for confirmatory testing. The present protocol does result in some of those with colour deficiencies, although in the main those with relatively minor levels of impairment, being accepted and there is at present no means of establishing whether their deficiency is of sufficient severity to be safety critical.

52. The HW-A lantern, using low intensity settings, could form a valid replacement for the HW-B and would be more compatible with the MCA Ishihara test protocol. However, as the same problems of supply and maintenance apply to both models, this is not a rational approach, except insofar as it enables the better validation data for the HW-A, given the consistency of rankings found between HW-A and HW-B lanterns (see paragraph 50), mean that HW-A data can validly be used to help determine the best option to adopt. The other available lanterns all have similar supply and maintenance problems and are less well validated that the HW-A or B, therefore they should not be adopted. The other established tests (paragraphs 30-40) that do not use coloured light lanterns are not recommended as they would give results that would differ widely from those using the current procedures. Of the three screen-based tests only the CAD is sufficiently well validated to be considered for use (paragraphs 41-46).

53. There are two benchmarks that could be used to determine a valid CAD threshold functionally equivalent to that in current use by MCA; i.e. that should ensure safety critical colour vision performance requirements are met and that unjustifiable discrimination by failing applicants capable of performing colour related safety-critical tasks is minimised:

- Age related equivalence:
  As seafarers with normal colour vision continue to perform lookout duties into their sixties, without apparent excess risk to maritime safety from colour perception, it is proposed that the threshold when CAD is used as a confirmatory test is set at the upper limit that defines normal red/green colour vision in those aged 65. Such information is available already for CAD: the upper, 65 years old limit of normal red/green colour vision (for binocular viewing) is 2.25 CAD units and that for yellow/blue colour vision is 2.44 CAD units.

- Holmes Wright A equivalence:
  The available data on CAD/HW-A equivalence forms the basis for this approach. As noted in paragraph 50, there is a constant relationship between the results of testing with HW-A and HW-B lanterns and so information from the more fully investigated HW-A lantern can be extrapolated to determine equivalent thresholds for the HW-B. This approach would have a number of benefits and would be consistent with approaches that are likely to be adopted in other transport sectors. A maximum CAD threshold limit of 2.35 would be roughly equivalent in outcome to the current MCA protocol for navigational lookouts (Appendix 1). All those with at binocular red/green CAD threshold of <2.35 CAD units pass the HW-A lantern (Figure 7). This threshold could also be regarded for all practical purposes as reflecting *'normal colour vision'.* All

normal trichromats pass with a threshold less than 2.35 units, while only ~ 4 to 5% of the mildest deutans also pass.

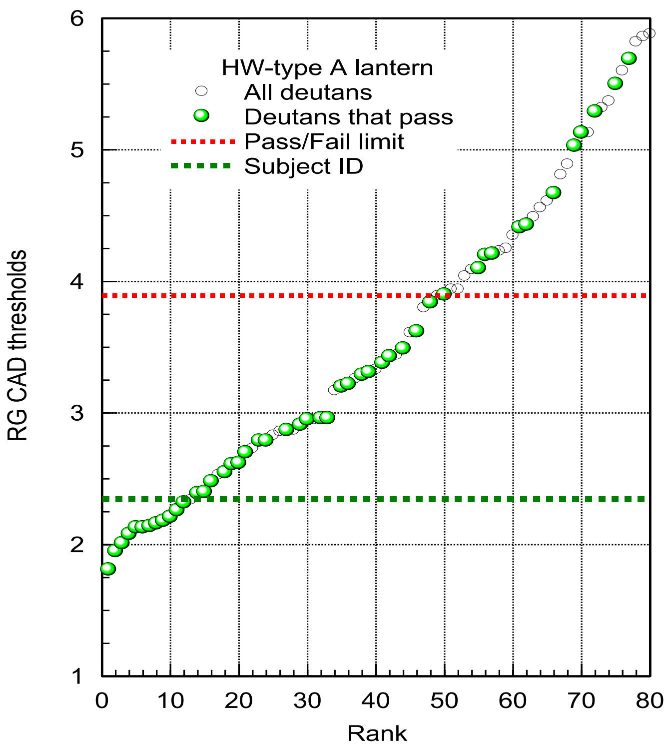


Figure 19. Enlargement from Figure 7 with proposed 'normal colour vision' (horizontal line of green dots) separating all those applicants that pass the HW-A protocol. The red dotted line defines the upper limit that is equivalent to the CAA HW-A protocol. The latter passes all normal trichromats and ~ 22% of deutan subjects and less than 1% of protans. The number of subjects that fail HW-A with a threshold less than 4 units equals those that pass with a threshold ≥ 4 units.

54. Although beyond the remit of this analysis, a similar approach could be applied to other groups of seafarers, such as engineers with visual tasks that are not at the limits of colour perception. An equivalent that passes as safe all those with thresholds < 4 (i.e., ~ 22% of deutans and significantly less than 1 % of protans)(Figure 4). This could be considered as a functionally adequate level of colour vision for less demanding tasks. All applicants with thresholds < 2.35 pass the HW-type A. The adoption of a functionally adequate category based on a CAD threshold of < 4 units is statistically equivalent to limits widely used in other transport sectors based on initial Ishihara screening with zero errors and a pass on HW-A lantern.

55. Given the identification of the same CAD threshold both from considerations of age related decline and from HW-A equivalence, 2.35 CAD units is an appropriate and justifiable threshold to adopt. For the HW-A equivalence approach CAD test thresholds are directly proportional to the cone photoreceptor signals generated and the upper-threshold, age-corrected limits for normal vision have been incorporated in the test. However it should be noted that any approach that simply substitutes one test for another or varies thresholds without a clear rationale is going to be less reliable than one that is based on a detailed visual task analysis for the duties being undertaken.

56. As the majority of colour deficiencies are congenital and relate to red/green discrimination it would be rational to put most effort into ensuring that a person's colour vision capability is carefully evaluated when they first start to work at sea. Periodic testing for congenital red/green defects in younger seafarers, the only sort identified by Ishihara plates, is unnecessary, and not entirely logical, despite it being required every six years in the STCW Convention. Acquired yellow /blue and red/green deficiency are commonly

secondary to other eye disease and so are rarely present early in life but are more frequent later in life. Yellow/blue testing has been seen as less directly safety critical, although this may no longer be valid if visual displays using a full range of colours are used for navigation, for instance to display charts. Any underlying disease can, however, also interfere with other aspects of vision that can be safety critical. The benefits of testing for acquired defects routinely has not been assessed here, but the general use of a test method such as CAD that could identify early stage retinal or systemic diseases that affect vision might be desirable.

57. This review is solely concerned with colour vision. In reality all aspects of visual function interact with each other to affect overall visual performance. For instance: dark adaptation, acuity under variable lighting conditions, glare and visual field integrity. A holistic view on the required visual performance 'envelope' is needed and, as for colour vision, this needs to be based on an analysis of visual task requirements coupled with information on the limits of visual performance in those without any impairment.

<u>Proposed options for revisions to MCA procedures</u>

58. The previous section identified CAD as the preferred test to substitute for the HW-B as a confirmatory test and identified an appropriate threshold to use. Any substitution needs to be of a single new test for the one it replaces. Acceptance of more than one confirmatory test will lead to lack of clarity in decision taking and to those wanting employment at sea seeking to perform multiple tests until they manage to pass one. The ease with which pass/fail CAD threshold limits can be revised in the light of new knowledge about safe levels of colour deficiency and the ability to go back over past records to identify anyone affected by a change to the threshold, the potential to detect yellow/blue defects associated with other eye pathology and the ability to use the same testing hardware to assess other facets of visual function would be additional advantages from adopting this test

59. As noted, the differences between HW-B and CAD mean that the CAD, or indeed any other test, cannot be direct like for like substitute. Because CAD has the advantage of determining the severity of colour deficiencies in both red/green and yellow/blue axes a quantitative threshold has to be established as the basis for taking pass/fail decisions. Based on the analysis of results and comparison with other tests and with previous practices, a threshold of 2.35 CAD units is recommended and can be justified.

60. Given that some of those assessed as fit will be close to the pass/fail criterion it may be appropriate to repeat CAD testing periodically, and give the person advance notice that their continued employment for lookout duties may be endangered if they fall below the threshold. Retesting would also provide an opportunity to assess whether any loss of yellow/blue sensitivity is also accompanied by yellow/blue loss, indicating the presence of early stage eye disease in those having repeat CAD tests. However this would only be applicable

to those who had the CAD as a confirmatory test because they had failed at the Ishihara.

61. Four options are proposed and it is for MCA to decide on the preferred one:

Option 1: continue with existing Ishihara protocol (>2 failed on 24 plate or >3 failed on 38 is unacceptable) and then use the CAD test with a threshold of 2.35 CAD units in a similar way to the current testing with HW-B lanterns.

This option would:

- Retain some of the limitations of the current MCA testing protocol. If ~ two or less IH errors are allowed on the first 25 plates of the 38 plate edition ~ 98% of normals pass, 4% of deutans pass and less than 1% of protans pass. The number of plates failed on the Ishihara test is not a good indicator of severity of colour vision loss (see Figure 5). Hence some of those with an unquantified level of colour vision deficiency are considered fit.

- Reduce the misclassification that currently arises from the use of the HW-B lantern where, when used strictly according to the MCA protocol, virtually everyone fails. Use of a CAD thresholds < 2.35 CAD units would ensure that all normals pass as well as 5% of the least affected deutans pass.

- Be expected to result in the number of requests for CAD tests being broadly similar to those for HW-B tests at present.

Option 2: revise Ishihara test protocol to minimise numbers of those with colour deficiencies being excluded from more detailed testing. Use zero errors on Ishihara as the criterion for confirmatory testing. Then test with CAD using threshold of <2.35 CAD units.

This option would:

- ensure that all those with colour vision deficiencies that are above the threshold of 2.35 CAD units would be correctly identified. It would also mean that some of those who do not reach this threshold can be advised that they may be suitable for other less colour vision demanding types of work at sea, as proposed in paragraph 51.
- result in increased demand for CAD tests, compared to option 1, but it would reduce the frequency of misclassification of those who are normal as colour vision deficient, as happens using HW-B. It would also further reduce the number of those incorrectly identified as normal because of the acceptance of up to two errors on IH. This option would also give more accurate results with potentially fairer and more widely applicable outcomes. It would be at variance with the first stage of the currently suspended IMO requirements for testing based on CIE 143:2001.

Option 3. One off use of CAD at start of career for all those with navigational lookout duties. A life-long certificate that indicates absence of congenital colour deficiencies is justified. At above 40 years of age periodic assessments could be justified to test for acquired deficiencies. Red/Green could be tested with Ishihara plates as is now done. Yellow/Blue would either require use of the CAD test or one of the other tests such as HRR that also assess this.

This option would:

- Align maritime practice with that in other sectors of transport where there are colour critical tasks.
- Mean that ADs had to test far fewer seafarers for colour vision using Ishihara.
- Create the need for CAD test facilities would need to be available to test c 500 officer cadets and a smaller but unquantifiable number of trainee deck ratings each year. For cadets this could be practically organised in the vicinity of training colleges, especially as a small number of ADs see the majority of cadet applicants. Tests for ratings could not be geographically localised so easily.
- Improve the reliability of the initial test, which determines career choice, by making it less dependent on ADs' adherence to Ishihara test requirement and greatly reducing the opportunities for cheating (see paragraph 33 and Figure 6).
- Pose practical problems about the need for six yearly testing to comply with international Convention requirements and raise complex issues about timing for repeat tests to detect significant deterioration from disease or aging.

Options 1-3 are suitable for meeting MCA's current needs. They vary in their implications for the Agency, for seafarers and for Approved Doctors.

Option 4: perform visual task analysis for present day navigational lookout duties and use the results to identify the most colour critical tasks. The pass/fail CAD threshold could then be based on the level of colour deficiency that is compatible with safe performance of these tasks. The analysis would identify the threshold at which performance of colour-critical tasks was the same as in those with normal trichromatic vision. The overall approach to testing could then be optimised using a mix of test methods, but with a better-validated threshold for unsafe levels.

This option would:

- Improve the validity of the colour vision testing process by providing an up to date evidence base for it. It may well enable more applicants with congenital colour vision deficiency to have maritime careers and would lead to validated pass/fail thresholds for deutan and protan like deficiency.

- Make existing procedures more secure by determining red/green and yellow/blue thresholds and the various categories of results that are appropriate for safe performance of navigational lookout duties. This information could form the basis for sound standard setting and determination of the optimal test protocols.

- Form a valid basis for achieving an international consensus on the colour vision test protocol to be adopted by IMO.

63. All options would mean that the three rooms in MCA Marine Offices dedicated to HW-B testing could be released and that staff time would not need to be spent on testing unless CAD was made available at MCA centres. There are already a number of locations where CAD tests are available in UK, some of these would be willing to perform the assessments required by MCA, usually on an item of service basis (Appendix 4). MCA would need to decide which of these centres to work with and agree terms. MCA will also need to take a view on whether any additional centres are required and, if so, how these should be funded.

64. The lantern tests are a long established part of the maritime world. Seafarers, employers and regulators will all need to be made aware of and understand that a test that is less directly related to the critical visual tasks of navigational lookout is capable of providing a fairer and more valid assessment of colour vision that existing protocols.

Recommendations

67. The following are our recommendations, for consideration by MCA:

1. MCA should move from the use of the HW-B lantern to the CAD test for confirmatory testing of colour vision in those with navigational lookout duties. Testing could either be arranged on an item of service basis from those who already have the test or by creation of dedicated MCA facilities. A programme of information and consultation on the proposed changes will need to be mounted.

2. Based on the available evidence, the preferred option at this stage is 2 (see paragraph 61). This would improve the ability of the two-stage process to reach valid decisions, with potential benefits for maritime safety. Compared to use of HW-B, it would reduce the numbers rejected from service who have normal colour vision. An additional advantage would be comparability with procedures in commercial aviation where there is already practical experience of the use of a similar protocol and where, in future, both maritime and aviation sectors could benefit by sharing developments in knowledge.

3. If option 1 was to be adopted there should be no reduction in the current standards of maritime safety as compared to the present position, and it would reduce unjustifiable discrimination. The costs of purchase of CAD tests could be offset against the savings in time and accommodation

needed to perform HW-B tests in MCA Marine Offices.

4. An intermediate approach between options 1 and 2 would be to require CAD testing on all those who made any errors on Ishihara testing at the start of their careers, thus increasing the validity of the test protocol at the time when this was critical to a person's career choices. The current Ishihara pass criterion could be used for subsequent tests.

5. Options 3 and 4 are rational approaches, but 3 would make MCA practice out of line with the currently suspended IMO Convention requirements and a lead to a considerable increase in CAD tests. This would be offset by a reduction in the requirements for ADs to routinely perform Ishihara tests. Option 4 would need a programme of research to determine visual task requirements. Because of their implications both of these options would need to be discussed internationally and in the case of option 4 would require agreement on support for the required investigations.

6. Ishihara testing is performed in a not entirely consistent way by MCA Approved Doctors (Paragraph 13), especially regarding the lighting used and the order and mode of display of plates: precision could be improved by specifying the testing protocol in more detail and taking steps to ensure that it is being followed.

7. The feasibility of developing a simplified low cost version of the CAD test suitable for standard IT equipment and used by ADs in the quick screening mode could also be explored.

Appendix 1

MCA colour vision testing requirements (extract of MSN 1839)

Deck officers and ratings - Colour vision should be tested by the Approved Doctor with Ishihara plates, using the introductory plate, and all the transformation and vanishing plates. Those used should be recorded on the medical report form (ENG 2). Candidates who fail the Ishihara colour plate test may apply to one of the MCA's nominated Marine Offices listed at Annex C to this MSN, for their colour vision to be re-tested using a Holmes Wright B lantern. (MSN 1839)

Supplementary guidance on colour vision testing (extract of ADG 14 2010 manual)

Colour vision

Initial testing

The AD must ensure that the seafarer meets the colour vision standards. To comply with international guidelines, testing for all seafarers should be done with the standard Ishihara plates. Some screen-based tests for colour vision are now available, however to ensure consistency, ADs should continue to use the book of Ishihara plates. Testing should be carried out at every medical examination unless the AD has their own record of a previous medical where the test has been passed within the previous four years. Illumination should be good north facing daylight or with daylight fluorescent lighting. Incandescent lighting is unsuitable because of its colour balance. The criteria for a pass are two or less misreadings on the 24 plate test, or three or less misreadings on the 38 plate test. It is essential that seafarers applying for certificates of competency as deck or dual career (merchant/fishing) officers have full colour vision.

When testing a seafarer for the first time, special care must be exercised to ensure that the test is properly conducted. Such testing should not be delegated, and the AD should be aware that those with problems have been known, on occasions, to memorise the sequence of Ishihara plates. An inappropriate pass causes major problems for the seafarer and their employer if detected at a subsequent medical.

The Ishihara test is an effective screening test but where supplementary testing, see below, is performed the results of the supplementary test will determine any restrictions to be placed on the seafarer.

Supplementary testing – Deck

A deck applicant who fails the Ishihara test may arrange for their colour vision to be re-tested free of charge, using the Holmes Wright B Lantern, at one of the 3 MCA Marine Offices (see annex B of MSN 1839) that offer lantern tests. The AD should withhold the issue of an ENG 1 until the test has been carried out. Failure in this test will mean that a medical certificate may only be issued with the restriction "not fit for lookout duties" (and/or solo navigational watch). Although there is a tick box on the ENG 1 form relating to fitness for lookout duties, non-fitness must also be written as a restriction on duties.

A seafarer who is referred for a lantern test should not be issued with an ENG 1 until the results of the lantern test have been returned to the AD. Alternatively, the AD can offer to issue the seafarer with an ENG 1 suitably restricted to 'no look out duties', 'daylight duties only' or 'no work with coloured cables' from the outset.

In cases where a seafarer being examined by a non-UK based AD fails Ishihara, the AD should advise the seafarer of their right to attend for a lantern test in the UK if they choose to although it should be pointed out that the likelihood of passing a lantern test is small. It may aid their decision on whether to travel for a lantern test if they have additional investigation by an optometrist or ophthalmologist locally to determine the severity of their colour impairment as it is very unlikely that anyone who has more than a minor degree of impairment will pass a lantern test. Unfortunately there are no acceptable equivalent lantern tests outside the UK.

Follow up

Where a seafarer has failed the Ishihara test but has subsequently passed a lantern, City University or Farnsworth D15 test, they should be issued with a note by the tester, on letter-headed paper, giving details, including the date and location of the test and the name of the tester. Presentation of this letter at subsequent medicals should generally obviate the need for repeat tests. Clear pass results should normally be considered valid for the duration of the seafarer's career.

Most colour vision defects will be found in new seafarers and appropriate vocational advice should be given. Cases do occur where defects are detected in seafarers who previously apparently met the standards. The AD should normally try to obtain details of past test results and contact the AD concerned. It may be appropriate to recommend that the seafarer seeks an ophthalmological opinion in case of any undetected eye disease.

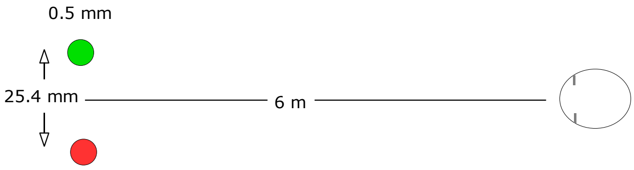
A seafarer with colour vision defects working or potentially working in both deck and engine departments should be tested and restricted, if necessary, in relation to both. For new entrants to officer cadetships where restricted duties are impractical because the full range of training cannot be carried out, those with defects should be made permanently unfit but, where appropriate, advised of the duties for which they could be suitable if they chose to apply for a different cadetship.

Appendix 2

<u>The Holmes Wright Lantern</u>

Pairs of coloured lights are displayed and the candidate has to identify them correctly. The lantern can be operated in two modes. In A mode the dimension of the aperture are wider and two levels of lighting can be used. The B mode has a smaller aperture and a single lighting level, which aims to mimic ship navigation lights at 2 nautical miles.

Stimulus dimensions employed in HW types A and B lanterns. The dimensions given are based on measurements carried out on available lanterns using a travelling microscope with a resolution of 0.01mm.



Traditional geometry for marine lantern tests (HW type B)

Demonstration colour uses 5.08 mm aperture size at 6m (~ 3 min arc). Simulation of marine red / green lights is achieved with two apertures (0.5 mm each ~ 0.3') separated by ~ 25.4 mm, giving a simulated angular separation of ~ 15 min arc. This is assumed to simulate the angular subtense of real signal lights at 2 miles (1 mile = 1.609 km). The real lights on the boat will be separated by ~ 13.6m. Also the diameter of each light will be 26.8 cm?. The average light intensity of the signal lights is ~ 12.7 ± 0.9 mcd (mean ± standard deviation).

HW lantern colour pairs (D = 5.08 mm; d = 0.5 mm)
Demonstration mode: uses 9 filter wheel positions

00	**W W**
15	**R G**
50	**W R**
17	**R G**
11	**G G**
53	**G R**
55	**R R**
01	**G W**
33	**G G**

Comparison with HW type A lantern

This lantern employs three different intensity settings: Demo, High and Low. The HW type B lantern employs a single intensity setting. Demonstration aperture diameter = 5 mm, Test aperture diameter is 1.52 mm (subtending an angle of ~ 0.87'). The two lights are separated by ~ 25.5mm giving an angular separation of ~ 15' at a viewing distance of 6m. Type A with the low intensity setting turns out to be almost equivalent to type B (but type A is almost always used with the high intensity setting). The vertical or horizontal orientation of the two signal lights is not likely to cause any significant difference in performance.

Appendix 3.

The features of the CAD test

The Colour Assessment and Diagnosis (CAD) test was developed at City University London (Barbur et al., 1994[29]; Rodriguez-Carmona et al., 2005[30]) and has been adapted for use in aviation with support from the Civil Aviation Authority (CAA, UK), and the Federal Aviation Administration (FAA, USA). The CAD test has a number of advantages which make it ideal for setting up standard protocols that are safe and reliable and can be can be implemented.

Key features

- Complete isolation of colour signals (i.e., in the absence of colour signals the applicant cannot do the test)
- The test uses an internationally recognized and reproducible system of colour representation (i.e., the CIE – (x,y) chromaticity chart)
- The test can be used to assess both the RG and the YB chromatic mechanisms
- The measured thresholds are directly proportional to the signals generated by the coloured stimulus in cone photoreceptors
- The test cannot be learned and the applicant cannot make use of other cues
- The equipment supplied for the test includes a 30 bit, stable visual display system and an accurate photometer that can be used periodically (e.g. six monthly) to check the calibration of the display and to recalibrate automatically, if necessary.
- The test relies on extensive studies designed to assess accurately the variability in RG and YB chromatic sensitivity in young, normal trichromats. The CAD unit for RG and YB colour vision is based on the median thresholds estimated in 333 young, normal trichromats. All thresholds are expressed in standard normal CAD units for ease of use and understanding. The test accurately classifies the applicants class or colour vision as normal trichromacy, congenital deficiency (i.e. deutan-, protan-, or tritan-defect) or acquired deficiency
- The test results are not affected significantly by pupil size, higher order ocular aberrations, small refractive errors, moderate levels of absorption of short –wavelength light by the lens and small variations in viewing distance.
- The test employs normal, upper-threshold age limits derived from extensive studies carried out in 393 subjects with normal colour vision (Barbur and Rodriguez-Carmona, 2015). Upper-thresholds limits for normal colour vision that describe monocular viewing are also provided.
- The CAD test provides 'Fast screening' option which uses the upper-threshold, normal age limits to screen for both RG and YB colour vision in less than one two minutes. Approximately 95% of applicants with normal colour vision pass. ~ 5% of normal trichromats and all subjects with congenital and acquired deficiency fail. Those that fail have the option to

carry out a Full CAD test which determines the applicant's class of colour vision (i.e., normal, deutan, protan or acquired) and the severity of RG and YB loss.

- Pass / Fail limits of colour vision loss that are considered safe within visually demanding working environments have been obtained for commercial pilots and TfL train drivers and the test is also being used with normal age-corrected, upper threshold limits to select normal trichromats within the fire service and for air traffic control.
- In addition to colour vision assessment, the Advanced Vision and Optometric Tests (AVOT) system (implemented on the same equipment that runs CAD) also offers tests to assess photopic and mesopic spatial vision, scattered light and rapid flicker and motion sensitivity.

There are over 120 systems available worldwide, some with several units at the same centre, and the number of users is increasing every year.

Appendix 4

Principle CAD users

Airlines / Aviation Authorities / Research Labs		Medical practitioners / Hospitals / Universities
Aeglia Medical Center, Netherlands		Anglia Ruskin University (UK)
Aeromedical Research Labs (USA)		Antwerp University Hospital (Belgium)
Bangkok Aviation Institute		Aston University, Optometry & Visual Science
Belgian Army		Belfast City Hospital, Northern Ireland
Cathay Pacific		Birmingham Eye Hospital (UK)
Civil Aviation Authority (UK)		Brookdale Medical Centre (UK)
Defence R&D Canada		Buskerud and Vestfold University, Kongsberg, N
Emirates Airlines		Calhoun Vision Inc
Etihad Airlines		Cardif University, Optometry & Visual Science
Federal Aviation Administration (USA)		Dublin Institute of Technology
Home Office (UK)		FJ Parkes Clinics, Dods St. Brunswick,Australia
Italian Air Force		KUPA Medical Center (Lagos, Nigeria)
Lufthansa		Liverpool Hope University (UK)
Maina Group, Mumbai		Mastricht University Hospital (Netherlands)
National Air Traffic Control Centre (UK)		Mends Specialist Hospital (Nigeria)
National Defence, Toronto, Canada		Ophthalmology Dept, University of Toronto
Naval Medical Research Unit (Dayton, USA)		Papworth Hospital, Cambridge
Norwegian Aviation Authority		Raffles Medical Group, Singapore
Royal Air Force, Henlow (UK)		Rohan Medical Centre, Mumbai
South African Civil Aviation Authority		Stansted Aviation Medical
Spanish Airforce, Madrid		The Livingstone Clinic, Sydney, Australia
Swedish Air Force		Universidad Federal do Pará, Brazil
Swiss Institute of Aviation		University of Auckland (NZ)
Thailand Aviation Authority (Bangkok)		University of Braga, Portugal
Transport Canada		University of Leipzig, Germany
Transport for London		University of New South Wales, Australia
William J. Hughes Technical Center (USA)		University of Pretoria (Ophthalmology)
Wright-Patterson AFB (USA)		Vision Test Australia Pty
		VU Medical Centre, Amsterdam (Netherlands)
		Wisconsin University (Ophthalmology)

References

[1] Sharpe, L. T., Stockman, A., Jagle, H., & Nathans, J. (1999). Opsin genes,

[2] Carter T. *Merchant Seamen's Health 1860-1960: Medicine, Technology, Shipowners and the State in Britain.* Woodbridge: Boydell Press, 2014. Pp 101-115.

[3] Topley H (1959) Sight testing for the merchant navy. *Br J Physiol. Opth.* 16:36-47.

[4] Squire, T. J., Rodriguez-Carmona, M., Evans, A. D. B., & Barbur, J. L. (2005). Color vision tests for aviation: comparison of the anomaloscope and three lantern types. *Aviat.Space Environ.Med.* **76**, 421-429.

[5] CIE (2001) CIE (*Commission Internationale de L'eclairage*). Colours of signal lights. Vienna Austria. In: Standard: S004/E-2001 Bureau Central de la CIE.

[6] CIE (*Commission Internationale de L'eclairage*) Technical Report 'International Recommendations for Colour Vision Requirements in Transport' 143: 2001.

[7] STCW Code table A-I/9: Minimum in-service eyesight standards for seafarers. *Guidelines on the medical examinations of seafarers*. International Labour Organization, Geneva 2013. p63 http://www.ilo.org/wcmsp5/groups/public/---ed_dialogue/---sector/documents/normativeinstrument/wcms_174794.pdf

[8] Test Methods of Colour Vision in Seafarers with Navigational Look-out duties. Expert Workshop Kobe 20-21st January 2014. http://www.imha.net/images/2014-01_ST-KOBE_Colour_Vision_Workshop_report_final.pdf

[9] Minimum Colour Vision Requirements for Professional Flight Crew. Recommendations for new colour vision standards. http://www.caa.co.uk/application.aspx?catid=33&pagetype=65&appid=11&mode=detail&id=3560

[10] Birch J. Assessment of Colour Visions Requirements for Seafarers and Testing Procedures. City University, London. Marine Safety Agency Project 414 8/1997.

[11] Seafarer Medical Examination System and Medical and Eyesight Standards: Application of the Merchant Shipping (Maritime Labour Convention) (Medical Certification) Regulations 2010. MSN 1839. Maritime and Coastguard Agency, 2009. pp35-37. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/281947/msn1839.pdf

[12] Approved Doctor's Manual: Seafarer Medical Examinations, Maritime and

Coastguard Agency, 2010. ADG 14: Vision, pp 234-236. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/285747/mca_doctors_manual__links_.pdf

[13] Work instruction seafarers colour vision test by examiners conducting lantern tests. Quality management system documentation. Document number MCA 710/001. Revision: 02 Date: 29.07.04. This has been supplemented by the additional instruction that if there were only two errors on the fifth random run the combinations failed could be re-tested.

[14] J. Neitz and M. Neitz. The genetics of normal and defective color vision. *Vision Res* 51 (7):633-651, 2011.

[15] Barbur J L, Rodriguez-Carmona M (2015). '*Colour Vision Changes in Normal Aging'* in: *Handbook of color psychology,* Andrew J. Elliot and Mark D. Fairchild, Editors. Cambridge University Press, 2015.

[16] Barbur JL, Rodriguez-Carmona ML in*: Handbook of color psychology,* Andrew J. Elliot and Mark D. Fairchild, Editors Cambridge University Press, 2015 (in press).

[17] Koenker R. A Note on Studentizing a Test for Heteroscedasticity. *Journal of Econometrics*, 1981;17: 107-12.

[18] 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.

[19] Squire TJ, Rodriguez-Carmona M, Evans ADB, Barbur JL (2005) Color vision tests for aviation: comparison of the anomaloscope and three lantern types. *Aviat Space Environ Med*; 76:422-429.

[20] J. L. Barbur and M. Rodriguez-Carmona. Variability in normal and defective colour vision: consequences for occupational environments. In: *Colour design*, edited by J. Best, Cambridge: Woodhead Publishing Limited, 2012, p. 24-82.

[21] Bailey KGH. *How consistent are the results of Occupational colour vision tests?* Dissertation, University of Birmingham, UK, 2013.

[22] Wright WD. *Researches on normal and defective colour vision,* London: Henry Kimpton, 1946.

[23] Barbur JL, Rodriguez-Carmona M, Harlow JA, Mancuso K, Neitz J, Neitz M (2008) A study of unusual Rayleigh matches in deutan deficiency. Vis Neurosci 25:507-516.

[24] CAD Test, City University London. www.city-occupational.co.uk

[25] Minimum Colour Vision Requirements for Professional Flight Crew. Recommendations for new colour vision standards. http://www.caa.co.uk/application.aspx?catid=33&pagetype=65&appid=11&mode=detail&id=3560

[26] Paramei GV, Oakley B. Variation of color discrimination across the life span. *J.Opt.Soc.Am.A Opt.Image Sci.Vis.* 31 (4): A375-A384, 2014.

[27] Cone Contrast Sensitivity Test, www.innovasysi.com

[28] Color DX Test, Konan, USA. http://colordx.com

[29] Barbur JL, Harlow AJ, Plant GT (1994) Insights into the different exploits of colour in the visual cortex. Proc R Soc Lond B 258:327-334.

[30] Rodriguez-Carmona ML, Harlow JA, Walker G, Barbur JL (2005) *The variability of normal trichromatic vision and the establishment of the "normal" range.* pp 979-982. Granada.