RESEARCH ARTICLE

Color Vision Tests for Aviation: Comparison of the Anomaloscope and Three Lantern Types

Theresa J. Squire, Marisa Rodriguez-Carmona, Anthony D. B. Evans, and John L. Barbur

Squire TJ, Rodriguez-Carmona M, Evans ADB, Barbur JL. *Color vision tests for aviation: comparison of the anomaloscope and three lantern types.* **Aviat Space Environ Med 2005; 76:421–9.**

Introduction: A comparison of the results obtained with the Nagel anomaloscope and the Holmes-Wright Type A, Spectrolux, and Beyne aviation color vision lanterns was undertaken. The Joint Aviation Requirements (JAR) specify pass/fail limits for these four secondary color vision tests and the Ishihara screening test. The results for individuals on all five tests were studied. ***Methods:*** The color vision of 55 color-vision deficient and 24 color-vision normal subjects, mostly applicant pilots, was assessed using a battery of tests, including the Ishihara plates, the Nagel anomaloscope, and three lanterns. The testing methods and characteristics of the lanterns and anomaloscope were compared. ***Results:*** Of the color-deficient applicants, only deuteranomalous trichromats passed more than one of the four secondary JAR tests, but a pass on one test did not reliably predict a pass on another test. Three out of nine protanomalous trichromats passed the Nagel anomaloscope but failed all three lantern tests. Of the normal trichromats, 12 failed the anomaloscope and 12 failed the Beyne lantern. ***Discussion:*** Variability in pass/fail results can be attributed to many factors apart from loss of chromatic sensitivity. Some normal trichromats can fail both the Ishihara screening and the secondary tests. The approved secondary test varies between countries and the outcome of regulatory assessment depends on the color vision test used. Since the flight safety consequences of the current situation cannot be ignored, the development of a less variable technique for color vision assessment that is accepted internationally, allied with a better understanding of color vision requirements, is needed.

Keywords: Nagel anomaloscope, Holmes-Wright Type A lantern, Spectrolux lantern, Beyne lantern, Ishihara plates, color vision, screening.

AVIATION COMMUNICATION technology has improved greatly over the last 30 yr, but color vision standards and tests are largely unchanged. Color is used on the airfield, in navigation lights, in the flight deck instrument panels, and for air traffic control applications. Improvements in color display technology have made it possible to use a wide range of color combinations to improve visual performance for many different tasks. In addition to increasing the conspicuity of objects (2), color is also used for warning signals and for color coding segmentation and grouping operations (23,24). Historically, it was assumed that color-deficient individuals would not make suitable professional pilots as they would be unable to discriminate the colors used in aviation, so would be unsafe. The difficulty of defining how different or deficient an individual's color vision can be without being unsafe has often been avoided by requiring applicants to have normal color vision. Although the appropriate use of color signals can undoubtedly improve visual performance, the benefits of requiring normal color vision in the aviation environment are difficult to assess, largely because of redundancy in the color-coding of visual information such as the use of flashing lights, audio cues, or text information. Also, luminance contrast, object shape, size, and location in the visual field are more often used to distinguish objects rather than, or in addition to, using specific colors. Concern has been expressed that, although pilots must be able to discriminate between certain colors because their work involves the recognition of various color codes in an evolving environment, it is not certain if the present requirements and testing procedures are appropriate for the tasks pilots carry out in their profession.

As part of a wider project initiated and financed by the UK Civil Aviation Authority for reviewing the minimum color vision requirements for professional flight crew, this study investigated the current methods of testing for color vision deficiency as required by the Joint Aviation Requirements (JAR). Currently, some applicants travel to different countries within the member states of the Joint Aviation Authorities (JAA) in search of a color vision test they can pass. This situation arises because different member states use different secondary tests and protocols. In order to compare the results from different tests, a group of color-deficient and color-normal subjects were tested on four alternative JAR recommended secondary tests.

Color Vision Deficiency

Color and luminance contrast contribute significantly to the perceived conspicuity of objects. Color-deficient observers vary greatly in their ability to discriminate

From the Applied Vision Research Centre, The Henry Wellcome Laboratories for Vision Sciences, City University, London, UK (T. J. Squire, M. Rodriguez-Carmona, J. L. Barbur); and the UK Civil Aviation Authority Safety Regulation Group, Aviation House, Gatwick Airport, UK (A. D. B. Evans).

This manuscript was received for review in September 2004. It was accepted for publication in January 2005.

Address reprint requests to: Prof. J. L. Barbur, Applied Vision Research Centre, The Henry Wellcome Laboratories for Vision Sciences, City University, Northampton Square, London EC1V 0HB, UK; j.l.barbur@city.ac.uk.

between different colors depending on the severity and type of their deficiency. Color vision deficiency is mostly congenital, and is fairly common, affecting almost 8% of men and 0.4% of women. Red/green deficiency is the most prevalent, and can vary from very mild, when color discrimination is not significantly different from that of color normals and is virtually unnoticeable in everyday life, to very severe, when visual performance can be significantly impaired. Congenital color vision deficiency does not affect other properties of vision such as acuity or contrast sensitivity. The type and severity of deficiency that a person inherits remains unchanged throughout life, although chromatic sensitivity can decrease gradually with age (21).

Humans with normal trichromatic color vision possess three distinct classes of cone photoreceptors (10,11,17,22) in the eye: short- (S); middle- (M); and long-wave (L) sensitive photopigments, often described as blue, green, and red cones, respectively (see **Fig. 1**). The majority of congenital color vision deficiencies arise as a result of changes in the properties of L- and M-cones, which cause specific patterns of color discrimination loss (see **Table I**). The most severe congenital deficiency (dichromacy) is caused when either L- or M-cones are missing. When both cone classes are present, but one class fails to produce normal responses to light in the corresponding wavelength range, the subject experiences reduced red-green discrimination (anomalous trichromacy). This is frequently the result of shifts in the wavelength of peak spectral sensitivity of either L- or M-cones, but other factors can also be involved (1). Depending on whether it is the L- or the M-cones that are affected, the deficiency is described as either protan (red) or deutan (green), respectively. In both cases, however, the same color channel is involved and the subject fails to discriminate a similar range of colors.

Unlike red/green deficiencies, congenital S-cone deficiency (tritan) is extremely rare and affects yellow/blue sensitivity. The large separation in spectral sensitivity of S- and M-cones also means that any small shifts in the wavelength of peak sensitivity of S-cones (see Fig. 1A) would not cause any significant changes in yellow/blue chromatic discrimination. Congenital tritanomalous deficiency is, therefore, more difficult to detect.

Defective color vision can also be acquired as a result of eye disease, or as a side effect of toxic poisoning or medication. Although current aviation tests are not designed to detect tritan or acquired color deficiency, this may require further consideration, given the more extensive use of colors in the modern aviation environment that require yellow/blue discrimination.

JAA Color Vision Testing

Color vision is currently tested within the JAA member states by screening with Ishihara pseudoisochromatic plates (Kanehara & Co. Ltd., Tokyo, Japan) and then, if the applicant fails this, testing with the Nagel anomaloscope (Schmidt and Haensch GmbH and Co., Berlin, Germany) or a lantern test, depending on the member state where tested. The lanterns are designed to provide an indication as to whether the aviation signal lights from the control tower, or aircraft navigation lights, would be discriminated correctly and were developed when red, green, and white lights were the main safety critical colors used in aviation. They do not identify the type or quantify the severity of color deficiency.

The Ishihara plates are used as a screening test for red/green color deficiency throughout the world and are used for this within the JAA; they have been shown to be efficient for that purpose (3,5,25). The plates do not diagnose the type or severity of color deficiency, but simply identify a subject as normal or as having red/green deficiency. Ishihara plates consist of a series of color-defined numbers embedded within different colored dots. The plates are designed so that grouping of dots by color causes a number to emerge from the background that can be recognized correctly by people with normal color vision, but in the absence of normal color signals all the dots appear 'falsely of the same color' (pseudoisochromatic). Therefore, color-deficient observers either fail to see the number altogether or make mistakes in recognizing it correctly. The JAR regulations state that: "The Ishihara test (24-plate version) is to be considered passed if the first 15 plates are identified without error, without uncertainty or hesitation (less than 3 s per plate)" (20).

The Nagel anomaloscope (4) is a relatively complex and expensive optical instrument and is not often chosen as a secondary test for aviation purposes, although it is generally accepted as the 'gold standard' in clinical practice for quantifying and discriminating between deutan and protan deficiency. This instrument is based on color matching. The examinee looks through a viewfinder to see a disk split horizontally into two half fields. The top half of the disk consists of red and green light that can be mixed in different proportions by turning a scaled knob, so creating a range of colors from 0 (green) to 73 (red). The bottom half of the disk is illuminated with a spectral yellow light (see Fig. 1A and B), the brightness of which can be altered by turning another scaled knob. The examinee matches the appearance of the two half fields in both color and brightness by altering the red/green mixture ratio and the brightness of the yellow field. All subjects match the two fields over a range of settings on the red/green scale, with normal trichromats achieving a match at a consistent mid-scale range of settings indicating normal functioning of green and red photoreceptors in the eye. Match positions on the scale outside the normal settings and the extent of the range determine the type of color deficiency and its severity. The JAR states: "Anomaloscopy (Nagel or equivalent). This test is considered passed if the color match is trichromatic and the matching range is 4 scale units or less. . . ." This is, therefore, the requirement that was applied to establish a pass or fail for each subject in this test.

A lantern test consists of a series of colored lights, or combinations of pairs of colored lights, being shown to the observer, who sits at a set distance from the lamp. The observer has to name the color of each light. Incorrect naming of some colors indicates color deficiency. These lights are designed to simulate signal lights used in aviation and their colors are chosen appropriately.

TABLE I. CLASSIFICATION AND PREVALENCE OF CONGENITAL COLOR DEFICIENCY.

Number of Functioning Cone Pigments	Type	Denomination and Cone Affected	Occurrence* as % of Population: Men	Occurrence* as % of Population: Women	Hue Discrimination
Two	Dichromat	a) Protanope (red)	1	0.02	Severely impaired
		b) Deuteranope (green)	1.1	0.01	
		c) Tritanope (blue)	0.002	0.001	
Three	Anomalous trichromat	a) Protanomalous (red)	1	0.02	Continuous range from slight to severe impairment
		b) Deuteranomalous (green)	4.9	0.38	
		c) Tritanomalous (blue)	Rare	Rare	

*From 19—estimates based on various surveys.

The three lanterns recommended by the JAA are the Holmes-Wright (Type A is used for aviation purposes) (Rayner, England) (8), the Spectrolux (Switzerland), and the Beyne (Luneau, France).

METHOD

Four secondary tests approved by the JAA were compared: the Nagel anomaloscope; and three lanterns—Holmes Wright Type A, Beyne, and Spectrolux. There were 79 subjects (71 men and 8 women—all female subjects were normal trichromats) who were tested and whose results were compared. The study protocol was approved by the Research and Ethical Committee at City University. Each subject provided informed consent before participating. The subjects were mostly potential pilots tested on behalf of the UK Civil Aviation Authority, either applying for a private pilot's license, holding a JAA Class 2 (private pilot) medical certificate and wanting color-deficient restrictions removed, or applying for a commercial pilot's license, requiring a JAA Class 1 medical certificate. The other subjects were randomly selected volunteers.

Each subject was first screened using the Ishihara pseudoisochromatic 38-plate test (the first 25 plates only, as the other plates are to test innumerate subjects). The first 15 plates of the 24-plate version is the screening test stipulated by JAR, with no errors as the pass criteria. These plates are included in the 38-plate version so results from both versions could be compared.

The Ishihara plates were viewed at about 70 cm (arm's length) distance using a MacBeth easel lamp (Kollmorgan Corp., Waltham, MA) for illumination. The book was placed in the tray beneath the lamp and the illumination, equivalent to Commission Internationale de l'Eclairage (CIE) Standard Illuminant C (representing average daylight), was incident at an angle of 45° to the plate surface. The examiner instructed the subject being tested to report what number they could see as the pages were turned, and warned that they might not see a number at all. The first introductory plate was used to demonstrate the visual task, as everyone should see this as a "12." A viewing time of about 3 s was allowed for each plate, with undue hesitation or uncertainty noted, as this is considered a criterion for not passing the screening according to the JAR.

If any mistake was made on the Ishihara plates, then a battery of clinical tests was carried out. These tests were: the American Optical Hardy, Rand, and Rittler plates (6; American Optical Company, New York, NY); Farnsworth D15 (Richmond Products, FL); City University test (Keeler Instruments, Windsor, UK) (7); and the Nagel anomaloscope (4), the results of all of which were used to define the type and severity of the red/green deficiency of the subject. If the Ishihara plates were passed then only the Nagel anomaloscope test was carried out to confirm normal color vision. The outcome of the test battery resulted in the classification of the following groups: 36 deuteranomalous trichromats, split into 'very slight,' 'slight,' 'slight/moderate,' 'moderate,' and 'severe' levels of deficiency; 5 deuteranopes; 9 protanomalous trichromats, split into 'slight,' 'moderate,' and 'severe' levels of deficiency; 5 protanopes; and 24 normal trichromats. After classification, all subjects were tested on the three lantern types.

For the Nagel anomaloscope, only the dominant eye was fully tested and the other was checked to confirm it had the same sensitivity as expected if the deficiency was congenital. The test was done in two stages. The subject was first familiarized with the instrument controls. The subject was then asked to alter both the control wheels until the two half fields looked exactly the same color and brightness. The subject was not asked to name the colors. A few matches were made, with the examiner 'spoiling' the match in between each attempt. About 10 s were allowed for each match and then, to minimize any chromatic after images, the subject looked away from the instrument into the dimly lit room.

The second stage of the test was to determine the limits of the matching range. The initial matches made by the subject in the first stage were used as a guide for the examiner to set the red/green mixture ratio near to the estimated limits of the range. The subject was instructed to adjust only the luminance of the lower yellow half-field and see if an exact 'match' in both color and brightness could be made with the set red/green mixture in the upper half-field. The examiner altered the ratio of the red/green mixture, alternating between near the upper and near the lower limits, until the limits of the matching range were established.

The match was then compared with the average nor-

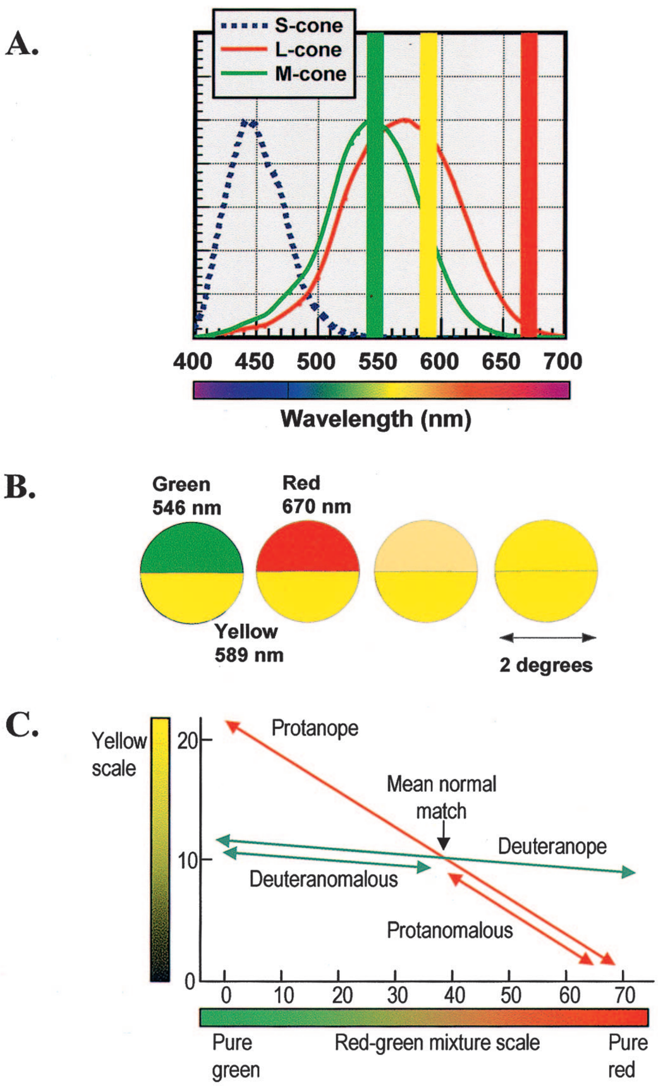

Fig. 1. A.) Graph showing the spectral sensitivity curves of the three types of cone photoreceptors that respond to the long (L-cones), middle (M-cones), and short (S-cones) wavelength ranges of the visual spectrum. The green (546 nm), yellow (598 nm), and red (670 nm) lights of the Nagel anomaloscope are superimposed. B.) Illustration of the Nagel anomaloscope split field. C.) Color matches obtained with the Nagel anomaloscope in protan and deutan deficiency. Protanopes and deuteranopes are able to match the yellow with all red/green mixture ratios by adjusting the luminance of the yellow comparison field. Protanomalous trichromats obtain matching ranges with an excess of red and deuteranomalous trichromats obtain matches with an excess of green in the matching field. The match made by normal trichomats is not usually accepted in anomalous trichromatism.

mal match on the Nagel anomaloscope given by 40 ± 4 scale units, calibrated for 68 color-normal subjects ± 2 SDs, with a yellow field value of 12 ± 2 scale units. If outside the normal settings, the type of deficiency was diagnosed by the position of the mean match, e.g., lower (more green) than the normal limits (0–35)—deuteranomalous trichromat; higher (more red) (45–73)—protanomalous trichromat. The matching range indicated the level of deficiency, with a larger range showing a more severe deficiency. The range is quoted as being two scale unit limits (e.g., 2–34 for a 'severe' deuteranomalous trichromat), and the extent of the range is the length of this range including both limits

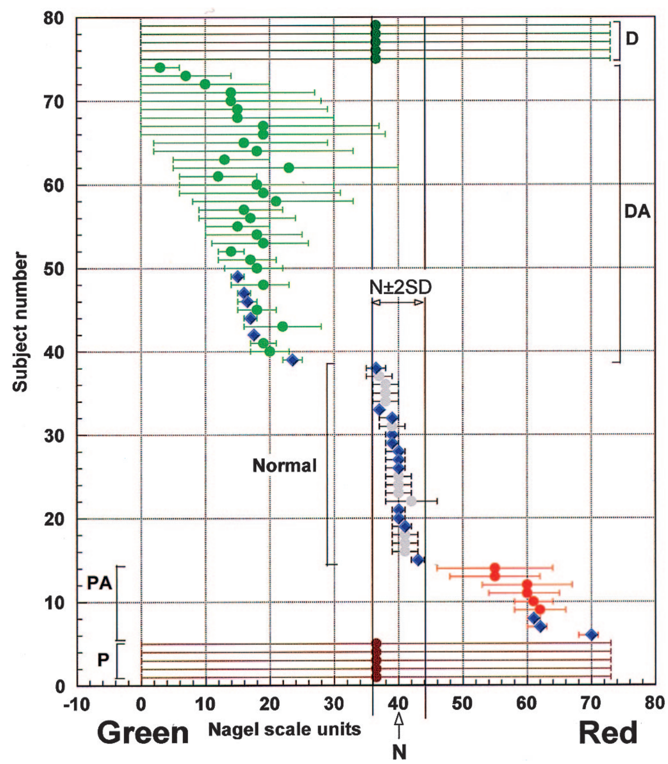

Fig. 2. Graph showing results of color matches on the Nagel anomaloscope. The solid, colored disks indicate the mean matching values on the red-green scale (blue diamonds indicate the mean matching values of those subjects who passed according to JAR; i.e., 4 scale units or less) and the horizontal lines depict the extent of the matching range. N indicates the value of the mean normal match. DA = deuteranomalous trichromats; D = deuteranopes; PA = protanomalous trichromats; P = protanopes.

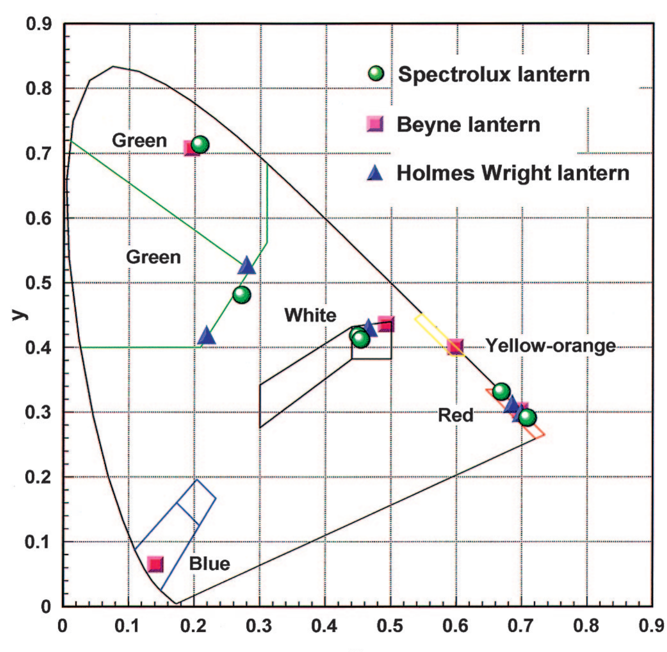

Fig. 3. CIE x,y chromaticity diagram (1931) showing the x,y chromaticity coordinates of the colors of the three different lanterns. The colored lines are the boundaries for the allowed chromaticity areas for the colors of light signals (13).

TABLE II. PROPERTIES OF LANTERNS.

Lantern [and Country of Origin]	Test Distance (m)	Colors of Lights	Number of Apertures	Aperture Sizes (min of arc)	Stimulus Duration (s)	Room Lighting Condition
Spectrolux [Switzerland]	2.5	RGW*	2, vertically separated	4.8	3	light
Beyne type 2 (now Tritest L3) [France]	5	WGRBY*	1	3	1	dim
Holmes-Wright Type A [UK]	6	RGW*	2, vertically separated	0.9	Max. 4 (manual)	normal/dark

*R = red; G = green; W = white; Y = yellow/orange; B = blue

(i.e., 33 in this example). The matching range for dichromats is the whole scale from 0 to 73, (matching yellow with all mixtures of red and green), with protanopes setting the yellow brightness scale very low (~0) for the red end of the range and high for the green end (~25). Deuteranopes set a consistent mid-scale brightness value (~12) for the whole range, as do normal trichromats (see Fig. 1C). This method was used for classification of each subject and for assessing a pass or fail by computing the red/green range of the match (**Fig. 2**). A JAR pass demands a range of 4 scale units or fewer on the Nagel.

The lantern tests were then carried out. The subject was seated at the appropriate distance (see **Table II**) from each lantern in turn. The test was explained and subjects were informed of the colors of the lights they would see and that they had to state which of these colors they saw after each presentation. The lights were always shown in the same sequence for each lantern and the tests were always carried out in the order Spectrolux, Beyne, and then Holmes-Wright Type A. The chromaticity coordinates and the luminance of every colored light on each lantern were measured with a telespectroradiometer (Model Minolta 1000, Konica Minolta, Buckinghamshire, UK). **Fig. 3** shows the chromaticity coordinates of each lantern plotted on a CIE 1931 (x,y) chromaticity graph (14). It can be seen that, although all three lanterns use red, green, and white lights, the specification of these colors is variable.

The Spectrolux lantern has twelve vertical pairs of combinations of red, green, or white lights, with two different intensities and hue of each color. The difference in luminance is relatively large in some of the pairs of lights, for example, white (413 cd · m^{-2}) and green (27 cd · m^{-2}), and two whites—the only large difference being in luminance (416 and 50 cd · m^{-2}). A total of 2 runs of the 12 pairs were carried out in the same order each time, and every color had to be named correctly (as 'red,' 'green,' or 'white') to pass. No demonstration of colors prior to the test was given, but subjects were advised of the colors they could expect.

The Tritest L3 lantern used in this study is a later version of the original 1950s Beyne lantern and will be referred to as a Beyne lantern throughout this paper. Five different colored lights, white, green, red, blue, and yellow/orange were individually shown to the subject, in that order, and each color had to be named correctly the first time in order to pass. No demonstration of the lights prior to the test was given, but subjects were advised of the colors they could expect. For this study, three runs of five lights were shown to each subject to determine if there was a learning effect with practice, but the pass/fail result was judged in accordance with aviation standards (the result of the first run). The five lights of the Beyne lantern also have large differences in luminance with blue being about 5 cd · m^{-2} and yellow-orange about 85 cd · m^{-2}.

The Holmes-Wright Type A lantern has a vertical pair of lights consisting of nine combinations with two different reds, two different greens, and a white light, all of similar intensity as seen by color normals. The subjects had to name what they saw as red, green, or white for each light. For this lantern three lights were shown to each subject before the test on the higher intensity 'DEM' setting (i.e., for demonstration only) and defined as 'red,' 'green,' and 'white.'

The test was carried out on the middle intensity 'HIGH' setting in a normally lit room, and if all nine pairs were named correctly, no more runs were carried out: the subject had passed. If any errors in naming were made then two more runs of the nine pairs were carried out. If there were no errors in these 18 pairs the subject had passed, but if any mistakes were made then the subject was dark adapted for 15 min and 1 more run of 9 pairs was undertaken in the dark. If the subject made no errors on this final run, this result was taken as a pass. If at any stage the subject called red 'green' or green 'red,' then the test was stopped and the subject failed.

RESULTS

The JAR procedure was followed to classify the results as pass or fail. All 55 color-deficient subjects failed the Ishihara plates by making at least 1 mistake in the first 15 plates of the 24-plate version, and so needed to pass 1 of the 4 JAA secondary tests in order to obtain an unrestricted Class 1 medical certificate. All the dichromats failed all the secondary tests; all the protanomalous trichromats failed all 3 lantern tests, but 3 passed the Nagel anomaloscope; and 14 of the 36 deuteranomalous trichromats passed at least 1 of the secondary tests.

Of the 24 normal trichromats that undertook the Ishihara plates test, 7 made between 1 and 3 mistakes in the first 15 plates. These seven would also go on to take the JAR secondary tests, with the possible risk of failing, even though their color vision was normal. The results show that all of those seven normal trichromats failed at least one of the secondary tests, apart from the Holmes-

TABLE III. THE PASS/FAIL RESULTS OF ALL THE SUBJECTS FOR EACH LANTERN TEST AND THE NAGEL ANOMALOSCOPE FOLLOWING THE JAR TESTING PROTOCOL.

Color Vision Type	Nagel Anomaloscope		Spectrolux		Beyne		Holmes-Wright Type A	
	Pass	Fail	Pass	Fail	Pass	Fail	Pass	Fail
DA[†] (N = 36)	6	30	3	33	8	28	6	30
D[†] (N = 5)	0	5	0	5	0	5	0	5
PA[†] (N = 9)	3	6	0	9	0	9	0	9
P[†] (N = 5)	0	5	0	5	0	5	0	5
Normal (N = 24)	12	12	21	3	12	12	24	0

[†]Color deficient subjects: DA = deuteranomalous trichromats; D = deuteranopes; PA = protanomalous trichromats; P = protanopes.

Wright lantern, which they all passed. **Table III** summarizes the results for all 79 subjects on all the JAR secondary tests.

Normal trichromats: There were 7 normal trichromats who failed the 24-plate Ishihara test. However, all the plates read incorrectly were considered by the examiner to be 'misreadings' (16), i.e., the unconscious 'filling-in' of gaps of a numeral (for example, a '5' may be interpreted as a '6', or '9' as an '8'), so these subjects would have passed a clinical test. Table III shows the results achieved by all the 24 normal trichromats on the secondary tests under JAR criterion. The Holmes-Wright Type A lantern was the only test with a 100% pass rate. About 14% failed the Spectrolux, calling a bright green 'white,' or a dull white 'red.' However, half of the color normals failed both the Nagel anomaloscope (match of 4 or more scale units) and the Beyne lantern (first run only), 92% of these calling the white 'yellow/orange' in the latter test. All but one of the seven subjects that failed the Ishihara plates and went on to do the further tests failed at least one of the secondary tests.

Fig. 2 displays the Nagel anomaloscope results, showing the extent of each matching range on the red/green scale and the mean of each range. The blue diamonds show the subjects that passed according to the JAR criterion (matching range of 4 scale units or fewer). It can be seen that only 12 out of the 24 subjects with normal color vision passed the Nagel test and of the 12 color normals that failed, 11 had a matching range of 5 scale units, only 1 greater than the permitted JAR range.

For the Beyne lantern test, as detailed in the Methods section, each of the single lights is shown just once (one run) to each subject for civil aviation purposes. However, for this project, each light was shown three times, following the same sequence each time. The results showed that 12 of the 24 normal trichromats passed Run 1; 2 of these subjects failed subsequent runs, and 5 subjects that failed Run 1 passed subsequent runs.

Dichromats: All 10 dichromats correctly identified a maximum of 3 out of the first 15 Ishihara plates (24-plate version), with all but 1 of the protanopes getting only 1 correct. All four secondary tests were failed by all the dichromats.

Protanomalous trichromats: All 9 protanomalous trichromats identified at most 4 Ishihara plates correctly, with one exception, an individual who identified 14 out of the 15 plates and then passed the Nagel anomaloscope. All the protanomalous trichromats failed all the lantern tests. However, three protanomalous trichromats passed the Nagel anomaloscope with matching ranges of 4 scale units or fewer. All protanomalous subjects failed the Beyne lantern test (Run 1 only), although three protanomalous trichromats passed one or both of the subsequent runs.

Deuteranomalous trichromats: The majority of the 36 deuteranomalous trichromats failed to read correctly 10 or more of the 15 Ishihara plates, even those with only a 'slight' deficiency. Of the 14 subjects with a 'slight' deficiency, 3 got 6 or fewer wrong. However, all of those with a 'very slight' deficiency got only 5 or fewer wrong.

Table IV analyzes the secondary test results from subjects with different severities of deuteranomaly. It shows that only 'slight' and 'very slight' deficiencies passed any of the secondary tests, with the exception of

TABLE IV. A COMPARISON OF THE NAGEL ANOMALOSCOPE AND LANTERN TEST RESULTS OF THE 36 DEUTERANOMALOUS TRICHROMATS.

No. of Subjects	Nagel	Lantern			Deficiency				
		Spectrolux	Beyne	H-W	v. slight	slight	slight/mod	moderate	severe
22	Fail	Fail	Fail	Fail		8	4	7	3
5	Fail	Fail	Pass	Fail		4		1	
2	Pass	Fail	Fail	Fail	1		1		
2	Pass	Fail	Fail	Pass	2				
1	Fail	Pass	Fail	Fail		1			
1	Fail	Fail	Fail	Pass		1			
1	Pass	Fail	Pass	Pass	1				
1	Fail	Pass	Pass	Pass	1				
1	Pass	Pass	Pass	Pass	1				

H-W = Holmes-Wright Type A.

TABLE V. COMPARISON OF THE RESULTS ON ALL SECONDARY TESTS OF FIVE SUBJECTS WITH SIMILAR NAGEL RANGES.

Subject No.	Severity of Def.*	Nagel Range	Nagel Result	Spect.†	Beyne	H-W
52	Slight	12–16	Fail	Fail	Fail	Fail
49	V slight‡	14–16	Pass	Pass	Pass	Pass
47	V slight‡	15–17	Pass	Fail	Fail	Fail
46	V slight‡	15–18	Pass	Fail	Fail	Pass
44	V slight‡	16–18	Pass	Fail	Fail	Pass

*Severity of deficiency; †Spectrolux; ‡Very slight; H-W = Holmes-Wright Type A.

a 'moderate' passing the Beyne lantern test, and one 'slight/moderate' passing the Nagel anomaloscope. Table III shows that a similar number of deuteranomalous trichromats passed the Nagel and each lantern, but Table IV shows in more detail that a pass on one test does not guarantee a pass on the others: only two subjects passed all three lanterns, one of which failed the Nagel, and one subject passed two of the three lanterns (the Beyne and the Holmes-Wright Type A) and the Nagel; these were all 'very slight' deuteranomalous trichromats.

Table V shows that deuteranomalous subjects who had similar Nagel results, in both size of ranges and scale unit values, achieved different lantern results. The only subject who failed the Nagel (subject 52) according to JAR also failed all three lantern tests. However, subject 47, who passed the Nagel with a range of only 3, also failed all three lanterns.

Table VI, detailing the different severities of deuteranomalous trichromats that passed various combinations of runs for the Beyne, shows that some of the deuteranomalous trichromats that passed the first run on the Beyne lantern (JAR pass criteria) then went on to fail one or more of the subsequent two runs. Similarly, some of those that failed Run 1 then passed one or more of the subsequent two runs.

Only 'very slight' and 'slight' color deficients passed the Beyne on the first run. Three out of six of the 'very slight' passed all three runs and the remainder failed all three. Of the four 'slights' that passed Run 1, two also passed the other two runs, but one failed both Run 2 and Run 3. It can be seen that one 'moderate' passed this lantern in both Run 1 and Run 3. This subject, in fact, also only made one error in Run 2, calling yellow/orange 'white.'

DISCUSSION

The results in this study show convincingly that the color vision tests currently approved in the JAR do not yield consistent results in passing and failing the same individuals. The Nagel anomaloscope is a completely different test from the lanterns (using color matching as opposed to color naming), but the colorimetric (Fig. 3) and physical characteristics of the lanterns (Table II) themselves are also quite different. The 24-plate Ishihara test results of 100% sensitivity and 70.8% specificity confirm the usefulness of Ishihara plates for screening red-green deficiency. Even in a larger study involving 471 color-normal subjects and 401 color-deficient subjects (identified with the Nagel anomaloscope using the clinical ophthalmologic criteria), the 16 most effective plates of the 38-plate Ishihara test were found by Birch (5) to have 98.7% sensitivity when 3 errors were allowed; however, the specificity was 94.1%. With the JAR pass criterion of not making any mistakes, virtually all red/green color-deficient subjects will invariably go on to take a secondary test. However, some normal trichromats will also need to take a secondary test.

The number of errors made on the Ishihara plates is not a useful guide to the severity of a color vision defect (12) except that it has been found that only color normals and a few mild anomalous trichromats make fewer than about 10 mistakes or misreadings on the 38-plate Ishihara test (5,9,16). The low specificity rate of the Ishihara plates is due to the design of the numerals that may lead to 'misreading' and should not be included as mistakes; i.e., as a failure of the plate. The interpretation of whether a wrong reading is a failure or a misreading often rests with the examiner. When the

TABLE VI. A COMPARISON OF THE RESULTS OF THREE CONSECUTIVE RUNS ON THE BEYNE LANTERN FOR DEUTERANOMALOUS TRICHROMATS.

No. of Subjects (N* = 36)	Beyne Lantern			Deficiency				
	Run 1	Run 2	Run 3	v. slight	slight	slight/mod	moderate	severe
5	Pass	Pass	Pass	3	2			
1	Pass	Pass	Fail		1			
1	Pass	Fail	Pass				1	
1	Pass	Fail	Fail		1			
1	Fail	Pass	Fail		1			
2	Fail	Pass	Pass		2			
4	Fail	Fail	Pass		3		1	
21	Fail	Fail	Fail	3	4	5	6	3

*N = Total number of subjects.

JAR are closely followed, then a noticeable 'uncertainty or hesitation' would lead to failure of that particular plate. As a result of applying this interpretation, 7 out of 24 of the subjects with normal color vision in this study failed the JAR Ishihara test.

Despite the procedure to take a secondary test, all but one of the seven normal trichromats that failed the Ishihara plates went on to fail one or more of these secondary tests and would, therefore, have been excluded from an aviation career as a pilot if they had been applying for this. Additionally, about 50% of the 24 color normals failed both the Beyne lantern and the Nagel anomaloscope (Table III). The failing of the Beyne lantern was largely due to the white light appearing rather yellow and being the first color shown, so the majority of normal trichromats that failed called this 'yellow/orange'; the Nagel anomaloscope was mostly failed because the restriction of the range for passing to 4 scale units or fewer is too harsh. The Holmes-Wright Type A lantern was the only test that passed 100% of the normal trichromats.

None of the tests passed subjects with the most severe deficiencies. Protans (L-cone deficient) are also assumed by some experts to potentially have more difficulties, as they perceive reddish colors as being dark. This is because neither of the other cone types are sensitive to the long wavelength region of the spectrum; this is not the case for deutans (M-cone deficient), as the L-cones are also sensitive in the middle wavelength range. This could lead to a pilot failing to observe a red obstruction light or a red navigation light sufficiently early to take appropriate avoiding action (even if they were able to recognize the color). On the other hand, perceiving red as a dark light might enable a protan individual to distinguish between red, and white or green, based on differences in brightness, even if the red color itself is not recognized. None of the lantern tests passed any subjects with a protan deficiency, but the Nagel anomaloscope did pass three 'slight' protanomalous trichromats. Fig. 2 shows that one of these three formed a match only at the extreme red end of the scale where the brightness of the match would also be very low. When three runs were carried out on the Beyne lantern, three protanomalous trichromats passed one or both of the subsequent runs. The reason for excluding all protanomalous subjects as professional pilots needs further investigation.

Our findings on the lantern tests (Table III) show that of the color-deficient subjects, only deuteranomalous trichromats are capable of passing any of three tests, and all of these, apart from one 'moderate,' had a 'slight' or 'very slight' deficiency. However, although the numbers that passed each test are not that different, it can be seen in Table IV that different subjects passed different lanterns. There were a total of only 14 out of the 36 deuteranomalous trichromats who passed any of the secondary tests. Of these 14 subjects, only 1 passed all 4 tests and 5 passed only the Beyne lantern. The severity of the deficiency also varied in those who passed any of the tests and did not seem to be consistent for each test.

There were six subjects with 'very slight' deuteranomalous deficiency whose color sensitivity was not particularly different from that of color-normal subjects, but even these subjects did not pass all the secondary tests, or even consistently the same tests.

The discrepancies in the results are partly attributable to the different testing criteria for each test. The colors used for the lanterns are clearly different, although the same names of 'red,' 'green,' and 'white' are used to describe them. These differences may not be important to those with normal color vision, but can make a major difference to the color discrimination of individuals with color deficiency if required to name these colors. The names of colors can be disputed even by color normals, particularly when presented in isolation against a dark surround. Color-deficient observers have reduced chromatic discrimination and it is, therefore, possible that the range of chromaticities they associate with the same color name is increased by comparison with normal trichromats.

The Holmes-Wright Type A and the Spectrolux lanterns both test the ability to name two colors when viewed simultaneously. The names given to the color of each light is often influenced by the colors present in the combination. The Beyne lantern only shows one light at a time, so any cue that may arise from comparison of color differences is eliminated. The Beyne and Spectrolux lanterns have lights with different luminance levels; perceived brightness differences can in principle be used to help identify selected colors. The fact that the Beyne lantern only shows each color once can also affect the outcome, since it is possible for color defective subjects to make some correct identifications by chance; the CIE (12) advises that at least five repetitions should be made to reduce chance probability.

Discussion with aviation color vision testing experts in different JAA member states has also produced useful insights into problems associated with the currently employed color vision tests. Both the tests and the protocols employed vary from state to state, even where the same test is used. In Germany, for example, the different colors of the lights of the Beyne lantern are demonstrated to the subjects, who are then tested on each color twice, rather than just once, and no mistakes are allowed (C. Stern. Personal communication; 2004). These differences, and other variations elsewhere, will affect the outcome of the tests. The JAA has chosen 4 tests in an attempt to standardize testing, but outside the 38 member states of the JAA there is an even wider range of tests in use (15,18,26).

CONCLUSIONS

This study has shown examples of the range of different color vision tests available to JAA member states. Comparison of these tests reveals the enormous variability and inconsistency of results, particularly when these are compared subject by subject rather than just examining the average pass/fail statistics of each lantern. There are several factors that can cause differences in color discrimination both within normal and color-deficient observers (12), and these factors can affect the outcome of the various tests in different ways (27). In addition, at least part of the large within-subject and intersubject variability observed in this study can be attributed to the differences in stimulus characteristics, testing methods, and the different instructions given to the subjects. The results also show that some normal trichromats that fail the initial screening may also fail the secondary test and that the outcome of regulatory assessment depends on which color vision test is used, which varies between countries.

Consistency is lacking in color vision testing and an aspiring professional pilot may be accepted without limitation in one country, and rejected outright in another. The different tests also reveal different aspects of color deficiency and the severity of outcome may or may not relate directly to the subject's ability to discriminate colors. The current situation is, therefore, unsatisfactory, both in terms of flight safety issues, as well as being potentially unfair to some applicants. A test that provides a quantifiable and accurate measure of color vision as a way of reliably grading the loss of color sensitivity would solve some of these problems. A better understanding of color vision requirements in the aviation environment is also needed so as to set sensible pass/fail criteria that are relevant and safe and can, therefore, be accepted internationally.

ACKNOWLEDGMENTS

Financial support for this study came from the United Kingdom Civil Aviation Authority.

We would like to gratefully acknowledge Dr. Hans Hafner and the Office Fédéral de l'Aviation Civile (OFAC), Switzerland, for kindly lending us their Spectrolux lantern so we could carry out this study. We thank Drs. Claudia Stern, Corinne Roumes, and Frans Larminier for useful discussions and information on color vision testing procedures within the JAR. We also wish to thank Jennifer Birch for her interest and advice with the interpretation of the color vision tests employed in this study.

REFERENCES

1. Barbur JL. Understanding colour. Trends in Cognitive Sciences 2003; 7:434–6.
2. Barbur JL, Forsyth PM. The effective contrast of coloured targets and its relation to visual search. In: Brogan D, Gale A, Carr K, eds. Visual search 2. London: Taylor & Francis Ltd.; 1992:319–28.
3. Belcher SJ, Greenshields KW, Wright WD. A color vision survey. Br J Ophthalmol 1958; 42:355–9.
4. Birch J. Diagnosis of defective color vision using the Nagel anomaloscope. Docum Ophthal Proc Series 1982; 33:231–5.
5. Birch J. Efficiency of the Ishihara plates for identifying red/green color deficiency. Ophthal Physiol Opt 1997; 17:403–8.
6. Birch J. Clinical use of the American Optical Company (Hardy, Rand and Rittler) pseudoisochromatic plates. Ophthal Physiol Opt 1997; 17:248–54.
7. Birch J. Clinical use of the City University test, 2nd ed. Ophthal Physiol Opt 1997; 17:466–72.
8. Birch J. Performance of red/green color deficient subjects on the Holmes-Wright lantern (Type A) in photopic viewing. Aviat Space Environ Med 1999; 70:897–901.
9. Birch J, McKeever LM. Survey of the accuracy of new pseudoisochromatic plates. Ophthal Physiol Opt 1993; 13:35–40.
10. Bowmaker JK, Dartnall HJA. Visual pigments of rods and cones in a human retina. J Physiol (London) 1980; 298:501–11.
11. Brown PK, Wald G. Visual pigments in single rods and cones of the human retina. Science 1964; 144:45–52.
12. CIE. International recommendations for color vision requirements for transport. Vienna, Austria: International Commission on Illumination; 2001. Report No.: 143.
13. CIE. Colours of light signals. Vienna, Austria: International Commission on Illumination; 2001. Standard: S004/E-2001.
14. CIE Proceedings. Cambridge University Press, Cambridge; 1932: 19.
15. Civil Aviation Safety Authority Australia. Designated aviation medical examiner's handbook. Version 2.2, section 2. Canberra, Australia: CASA; 2000.
16. Cole BL. Misuse of the Ishihara test for colour blindness. Br J Physiol Opt 1963; 20:113–8.
17. Dartnall HJA, Bowmaker JK, Mollon JD. Human visual pigments: microspectrophotometric results from the eyes of seven persons. Proc R Soc Lond B Biol Sci 1983; 220:115–30.
18. Federal Aviation Authority, Office for Aerospace Medicine. Version III of the guide for aviation medical examiners. Retrieved November 4, 2004, from: http://www.faa.gov/avr/aam/Game/Version_2/03amemanual/home/home.html
19. Hunt RWG. Measuring colour. 3rd ed. London: Fountain Press; 1998.
20. Joint Aviation Authorities. JAR-FCL 3. Flight crew licensing (medical). Section 1, subpart B. Class 1 medical requirements (amendment 2). Hoofddorp, The Netherlands: JAA; 2002.
21. Knoblauch K, Vital-Durand F, Barbur JL. Variation of chromatic sensitivity across the life span. Vision Res 2001; 41:23–36.
22. Marks WB, Dobelle WH, MacNichol EF. Visual pigments of single primate cones. Science 1964; 143:1181–3.
23. McIlhagga WH, Hine T, Cole GR, Snyder AW. Texture segregation with luminance and chromatic contrast. Vision Res 1990; 30:489–95.
24. McIlhagga WH, Mullen KT. Contour integration with color and luminance contrast. Vision Res 1996; 36:1265–79.
25. Sloan LL, Habel A. Tests for color deficiency based on the P.I.C. principle–a comparative study. Arch Ophthalmol 1956; 55:229–39.
26. Transport Canada. Canadian aviation regulations. Handbook for civil aviation medical examiners, subpart 4. Medical requirements. Ottawa, Ontario, Canada: Transport Canada; 2003.
27. Wright WD. Researches on normal and defective colour vision. London: Henry Kimpton; 1946.