Color Vision Tests in Pilots' Medical Assessments

Marie Marechal; Maxime Delbarre; Joachim Tesson; Carole Lacambre; Hélène Lefebvre; Françoise Froussart-Maille

INTRODUCTION: The aim of this study was to evaluate the ability of eight color vision tests to screen for and accurately measure hereditary color-deficiency in order to improve color vision assessment methods for aircraft pilots.

METHODS: This prospective study included 29 color-deficient subjects and 23 healthy subjects. All performed the following tests: Ishihara plates, Farnsworth D15, Lanthony desaturated 15 Hue, Munsell 100 Hue, Beyne and Fletcher-Evans CAM lanterns, Nagel anomaloscope, and the Color Assessment and Diagnosis (CAD) test. The sensitivity and specificity of color-deficiency diagnosis were evaluated for each test, as well as the test's relevance for assessing aircraft pilots.

RESULTS: The Ishihara plate test demonstrated a sensitivity of 0.97 and a specificity of 1.00 for color-deficiency screening. The CAD test and anomaloscope showed both a sensitivity and specificity of 1.00. The Beyne lantern, Fletcher lantern, Farnsworth D15, and the Lanthony 15 Hue tests all showed a specificity of 1.00 and sensitivities of, respectively, 0.69, 0.97, 0.58, and 0.79. During aircraft pilot selection tests, the CAD test classified 10% of color-deficient subjects as safe to fly, the anomaloscope 17%, and the Beyne and Fletcher lantern tests, respectively, 31% and 3%.

DISCUSSION: The discrepancy in results confirms that current color vision test protocols need to be reassessed. The CAD test could be an interesting alternative to the series of tests used to assess flight crew, but it seems more selective than current tests.

KEYWORDS: CAD test, anomaloscope, lantern tests, Ishihara test, color-vision deficiency.

Marechal M, Delbarre M, Tesson J, Lacambre C, Lefebvre H, Froussart-Maille F. *Color vision tests in pilots' medical assessments*. Aerosp Med Hum Perform. 2018; 89(8):737–743.

Color deficiency affects 9% of all men and 0.5% of women, many of whom are not aware they are colorblind.[5,9,13] In France, 1.5–3% of candidate French Air Force pilots show impaired color perception of which they are unaware.[16]

The United Kingdom (UK) Civil Aviation Authority (CAA) published new recommendations in 2009 about color vision for aircrew.[4] According to this text, access to information is first assigned to a symbol, text, or sound, but some elements of security remain conditioned by color, especially the Precision Approach Path Indicator (PAPI). Color is more and more present in cockpits and helps to rapidly alert the crew in case of emergency. Many colors are used, particularly in "glass cockpits." Such cockpits feature electronic flight instrument displays, typically on large multicolored liquid crystal display screens. Manufacturers are free to choose the colors displayed, but they must be readily discriminable by color-vision normal subjects.

In France, medical fitness requirements for pilots are defined by the 27 January 2005 decree and follow European Aviation Safety Agency (EASA) regulations.[1,2] For class 1 medical certificates (professional pilots), normal color vision is defined as the ability to pass the Ishihara 24-plate test (first 15 plates identified without error or hesitation) or to show normal trichromatic vision using the Nagel anomaloscope. For subjects who fail the Ishihara test, color vision is considered safe when candidates pass the Beyne lantern test without mistake or hesitation (colored lights switched on for 1 s at a distance of 5 m with 3 min of arc aperture) or if the matching range is 4 scale units or less using the Nagel anomaloscope.

A number of secondary tests are used and vary from country to country. Such variability in testing methods encourages "aeromedical tourism," with pilot candidates seeking aeromedical examiners in countries where color assessment standards are less demanding.[22] It now seems essential to standardize regulations worldwide to facilitate appropriate and fair methods for selecting flight crew.

From Percy Military Hospital, Clamart, France, and the French Military Health Services Academy, Paris. France.

This manuscript was received for review in October 2017. It was accepted for publication in April 2018.

Address correspondence to: Marie Marechal, 68 rue Fondary, 75015 Paris, France; marechal.m@live.fr.

DOI: https://doi.org/10.3357/AMHP.5009.2018

The Beyne, Holmes-Wright, and Farnsworth lanterns are doomed to disappear in the close future as they are no longer marketed. The Fletcher-Evans CAM lantern is still available for purchase, but its use is currently being discussed. The anomaloscope procedure is more difficult to implement as it needs to be performed by a highly experienced examiner.

In 2001, the North Atlantic Treaty Organization task force issued recommendations on the use of color in aviation in order to propose an international evaluation standard for color vision assessment.[17] They recommended routine screening for red-green, but also blue-yellow deficiency. Likewise, the recommendations issued by the UK CAA question the reliability of lantern tests and advise to use the Color Assessment and Diagnosis (CAD) test for civil aviation. The aim of the present study was to assess the performance of eight color vision tests, including the CAD test, to screen for and evaluate red-green hereditary deficiency in order to improve and to adapt current color vision assessment methods.

METHODS

Subjects

This prospective study was conducted between September 2016 and May 2017 in the "Centre Principal d'Expertise Medicale du Personnel Navigant" of Percy Military Hospital in Clamart, France. Each subject provided informed consent before participating. The protocol met the ethical principles for medical research of the Declaration of Helsinki.

Inclusion criteria were: on the one hand, military or civil subjects referred for color vision exploration as part of their selection, and on the other hand, healthy volunteers. All subjects underwent ophthalmological examination with measurement of their visual acuity. All subjects had a best-corrected visual acuity at 6/6. Subjects with any ophthalmological condition were excluded. All tests were conducted without sunglasses or tinted contact lenses.

Categorization was defined in two steps by an expert ophthalmologist. Firstly, subjects were classified into two groups: color-vision normal (CVN) and color-vision deficient (CVD) via all test results. All tests were used because of the risk of false negatives. The anomaloscope is often the reference test and was the most important for group classification, but a patient with a normal anomaloscope and other abnormal tests was classified as CVD. Secondly, CVD subjects were classified into four types (protanomalous, protanope, deuteranoumalous, and deuteranope) according the anomaloscope results.

Materials

All subjects were examined under the conditions defined for each test, with a break (5 to 10 min) between each test. Tests were performed with binocular vision, except the anomaloscope, for which the dominant eye was used. The color arrangement tests and the Ishihara plate test were lit with a fluorescent tube lamp having a 95 Color Rendering Index, a color temperature of 6500 K, and providing an illuminance of 300–400 lx. Subjects first performed the tests of the current color vision assessment protocol: the Ishihara plate and the Beyne lantern tests. They then performed the other tests in an aleatory order to limit the tiredness effect.

For the Ishihara 38-plate test, the first 25 plates were presented out of order, at 70 cm of distance, at a 45° angle. The subject had 3 s to answer for each plate. Based on the current aviation protocol, the subject was considered to pass the test if the first 17 plates were read without error.

Candidates were then assessed using the Beyne lantern test for aviation. Any hesitation or false answer, even immediately corrected, was considered as an error. The five different aviation lights (red, green, blue, yellow-orange, and white) were presented in random order.

In a first step, two series of lights were shown without naming colors and were not repeated in case of failure: each light shown for 1 s with 4 min of arc aperture (protocol 1 s / 4') and then each light shown for 1/25th s with 2 min of arc aperture (protocol 1/25th s / 2'). Subjects then underwent the French civil aviation test protocol in which each light is shown for 1 s with 3 min of arc aperture (noted 1 s / 3'). Candidates were informed of the light colors (red, green, blue, yellow-orange, off-white) but not when the lights were on. Two additional runs were performed if the candidate failed the first run. The test was considered a pass if the five lights were recognized either during the first run, or during two of the three runs. We evaluated the results after the initial run and after all three runs (noted 1 s / 3' x 3 runs).

The Fletcher-Evans CAM lantern test for aviation was presented for 2 s (determined manually) with an arc aperture of 0.9 min. If the subject made an error on the first run, the combinations the subject failed were retested. The test was considered a pass if all the combinations were correctly perceived at the end of the retests. We evaluated the results after the initial run and after the retests.

The Farnsworth D15 test was successful in the absence of confusion lines (circular scheme). We tolerated a confusion line between tiles 7 and 15 and tile inversions. The Lanthony desaturated 15 Hue test was successful if the scheme provided by the subject contained less than 2 confusion lines (the presence of 2 confusion lines was tolerated in patients over 40 yr of age). For both tests, the type of defect was determined based on alignment of the greatest number of confusion lines with the protan, deutan, and tritan axes.[15]

The Farnsworth-Munsell 100 Hue was then performed. The test was considered a pass if the subject's score was lower than his physiological score (age $\times$ 2 + 30), and the type of color vision deficiency (axis) was determined from the peak plot of the highest color error spike.

The anomaloscope IF2 (All-Color Anomaloscope, Tomey, Japan) with the Rayleigh equation was used. The test was performed by a trained examiner on the dominant eye in mesopic conditions. The test was first performed in automatic mode, then in manual mode (to accurately determine matching range boundaries). According to current aviation standards, color vision is considered safe if the matching range is less than

4 scale units from the Nagel anomaloscope. The Tomey anomaloscope uses a scale of 100 units (compared with 73 units for the Nagel anomaloscope). By converting the units, we accepted a matching range of 5 units with the Tomey anomaloscope.

The CAD test was performed in a mesopic condition as a three-step procedure. The subject was asked to indicate the direction of the colored stimulus movement using a joystick. First, in “learning mode,” the subject’s understanding of the test procedure was checked. In a second step, the “fast screening” mode was used to classify a vast majority of healthy subjects as “color-vision normal.” Finally, the “definitive CAD” step was performed if too many errors were observed in the fast screening step. The test determines the chromatic sensitivity level by presenting stimuli of varying intensity for each wavelength. This allows the examiner to determine the deficiency axis and severity according to a score (RG for red-green and YB for yellow-blue) in Standard Normal (SN) units.

The Cone Contrast Test, which is another existing electronic color vision test, was not evaluated here because it is not yet available in France.[19]

Statistical Analysis

Sensitivity, specificity, and positive and negative predictive values for the color-deficiency detection were calculated for each test. For quantitative assessments such as the anomaloscope and CAD tests, the diagnosis of color deficiency or normal color vision was used, regardless of the severity of the deficiency.

The relevance of each test for the selection of aircraft pilots (for class 1 medical certificates) was evaluated: the anomaloscope was considered successful if the matching range was 4 equivalent Nagel units or less, and the CAD test was considered successful for scores of 12 SN or less for protan subjects and 6 SN or less for deutan subjects. For each test, we evaluated the total number of tests and the success ratio, depending on the type of color-deficiency.

The qualitative and quantitative diagnostic outcome of each test was compared with anomaloscope results (gold standard). We evaluated the agreement of the number and proportion of CVD subjects diagnosed as having protan or deutan deficiency with anomaloscope results. For quantitative tests, we plotted receiver operating characteristic (ROC) curves for the diagnosis of dichromatism in the CVD population.

The medical statistics computer software SPSS® (IBM, Armonk, NY) was used for statistical analysis and ROC curves.

RESULTS

For this study, 55 subjects were recruited to participate. A total of 32 were classified in the CVD group and 23 in the CVN group. Three CVD subjects were excluded for missing results. No CVN subjects were excluded.

The mean age was 23 yr ± 6.09 in the CVD group and 26 ± 6.1 yr in the CVN group ($P = 0.075$). All subjects in the CVD group were men, whereas there were 9 women (39%) in the CVN group. In the CVD group, there were 11 deuteranomalous trichromats (37.9%), 7 protanomalous trichromats (24.1%), 6 protanopes (20.69%), and 5 deuteranopes (17.24%).

The sensitivity and specificity results, as well as the positive and negative predictive values are provided in **Table I**. All subjects of the CVN group were found to be fit for flight according to the anomaloscope and the UK CAD test. All Lanthony 15 Hue desaturated and Farnsworth D15 tests were normal. One CVN subject had a Farnsworth-Munsell 100 Hue score of 77, slightly greater than his/her physiological score of 72, but passed all the other tests. More selective results were obtained with the 1/25th s / 2' protocol of the Beyne lantern test with 11 failures for 23 tests (47.8%). One subject also failed the 1 s / 4' protocol. All subjects passed the 1 s / 3' protocol on the initial run. For the Fletcher lantern test, five CVN subjects made a mistake during the first run, but corrected their mistake during the first retest.

For CVD subjects, **Table II** shows the total number and success ratio for class 1 medical certificates, depending on the type of color vision deficiency. The results for each individual CVD subject are provided in **Table III**.

Only three tests assess the severity of the chromatic deficiency with a score: the anomaloscope, the CAD test, and the Farnsworth-Munsell 100 Hue test. **Fig. 1** shows the ROC curves for three parameters used in the diagnosis of dichromatism using the anomaloscope as the gold standard:

- Score RG of the CAD test;
- Difference between the physiological score and the score achieved for the 100 Hue test; and
- Score achieved for the Farnsworth-Munsell 100 Hue.

The areas under the ROC curves were 0.98 (0.94; 1.01) for the CAD test, 0.69 (0.52; 0.86) for the 100 Hue test when considering the score alone, and 0.65 (0.47; 0.83) for the 100 Hue test when considering the difference between the achieved score and the physiological score. The difference between the CAD test and the 100 Hue test was statistically significant ($P = 0.02$).

Table I. Sensitivity, Specificity and Predictive Values for Each Color-Deficiency Detection Test.

	Se*	Sp†	PPV‡	NPV§
Ishihara	0.97	1.00	1.00	0.96
Farnsworth D15	0.58	1.00	1.00	0.64
Lanthony 15 Hue	0.79	1.00	1.00	0.79
100 Hue	0.79	0.96	0.96	0.79
Beyne lantern				
1 s / 4'	0.79	0.96	0.96	0.79
1/25th s / 2'	0.97	0.57	0.76	0.93
1 s / 3'	0.76	0.96	0.96	0.76
1 s / 3' x 3 series	0.69	1.00	1.00	0.72
Fletcher Lantern				
1 presentation	1.00	0.78	0.85	1.00
2 retests	0.97	1.00	1.00	0.96
CAD test	1.00	1.00	1.00	1.00
Anomaloscope				
Automatic	0.97	0.96	0.97	0.96
Manual	1.00	1.00	1.00	1.00

* Se: sensitivity; †Sp: specificity; ‡PPV: positive predictive value; §NPV: negative predictive value; CAD: color assessment and diagnosis.

Table II. Based on the Results for the Different Color-Deficiency Tests: Number and Rate of Success for a Class 1 Medical Certificate.

	DEUTERANOMALOUS TRICHROMATS		DEUTERANOPES		PROTANOMALOUS TRICHROMATS		PROTANOPES	
TEST	*N* = 11	IN %	*N* = 5	IN %	*N* = 7	IN %	*N* = 6	IN %
Ishihara	1	9	0	0	0	0	0	0
Farnsworth D15	5	45	0	0	5	71	2	33
Lanthony D15	4	36	0	0	2	29	0	0
Munsell 100 Hue	3	27	0	0	2	29	1	17
Beyne lantern								
1 s / 4'	5	45	0	0	1	14	0	0
1/25th s / 2'	1	9	0	0	0	0	0	0
1 s / 3'	3	27	1	20	2	29	1	17
1 s / 3' x 3	6	50	0	0	2	29	1	17
Fletcher lantern								
1 presentation	0	0	0	0	0	0	0	0
2 retests	1	9	0	0	0	0	0	0
CAD test	2	18	0	0	1	14	0	0
Anomaloscope	4	36	0	0	2	29	0	0

The best cut-off levels to distinguish dichromats from anomalous trichromats are:

- RG >18.5 for the CAD test: sensitivity of 0.90 and specificity of 0.89;
- Difference between score achieved and physiological score >22 for the 100 Hue test: sensitivity of 0.80 and specificity of 0.53; and
- Score >112 for the 100 Hue test: sensitivity of 0.70 and specificity of 0.68.

Apart from the lantern tests, all tests provided qualitative diagnostic information on the color-deficiency axis. We compared the axis identified using the Ishihara plates, Farnsworth D15, Lanthony 15 Hue, Farnsworth-Munsell 100 Hue, and CAD tests with the diagnosis obtained using the anomaloscope.

Only the CAD test found the expected diagnosis for each subject. Qualitative diagnosis could not be established using Ishihara plates and the Farnsworth D15 test for, respectively, 19% and 41% of CVD subjects (mostly subjects with a low to moderate deficiency, with arrangement tests leading to "low discrimination"). When an axis was determined, it was always in line with that found using the anomaloscope. Three dichromats could not be classified as protans or deutans by at least one test. Axis-based diagnosis was incorrect for two deuteranomalous trichromats using the Lanthony D15 test and for one protanomalous trichromat using the Farnsworth-Munsell 100 Hue test.

Table III. Results of Pilots' Requirements for Each Subject of CVD Group According the Different Tests.

CVD					BEYNE LANTERN				FLETCHER			
(SUBJECT NUMBER)	ISHIHARA	D15	15 HUE LANTHONY	100 HUE	1/4	1/25TH/2	1/3	3 TESTS	1 TEST	2 RETESTS	CAD TEST	ANOMALOSCOPE
2	Fail	Pass	Pass	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Pass	Pass
4	Fail	Pass	Pass	Fail	Fail	Fail	Pass	Pass	Fail	Fail	Fail	Pass
5	Fail	Pass	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail
6	Fail	Pass	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail
7	Fail	Fail	Fail	Fail	Fail	Fail	Pass	Pass	Fail	Fail	Fail	Fail
10	Fail	Fail	Fail	Fail	Fail	Fail	Pass	Fail	Fail	Fail	Fail	Fail
12	Fail	Fail	Fail	Fail	Pass	Fail	Fail	Fail	Fail	Fail	Fail	Pass
13	Fail	Pass	Fail	Fail	Pass	Fail	Fail	Fail	Fail	Fail	Fail	Fail
14	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Pass	Fail	Fail	Fail	Fail
15	Fail	Pass	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail
18	Fail	Pass	Pass	Fail	Fail	Fail	Fail	Pass	Fail	Fail	Pass	Fail
20	Fail	Fail	Fail	Pass	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail
22	Fail	Pass	Pass	Fail	Pass	Pass	Pass	Pass	Fail	Fail	Fail	Pass
23	Fail	Fail	Fail	Fail	Fail	Fail	Pass	Pass	Fail	Fail	Fail	Fail
24	Fail	Pass	Fail	Pass	Fail	Fail	Pass	Pass	Fail	Fail	Fail	Fail
25	Fail	Fail	Fail	Fail	Pass	Fail	Fail	Fail	Fail	Fail	Fail	Fail
27	Fail	Pass	Fail	Pass	Pass	Fail	Fail	Fail	Fail	Fail	Fail	Pass
28	Pass	Pass	Pass	Pass	Fail	Fail	Fail	Pass	Fail	Pass	Pass	Pass
29	Fail	Fail	Pass	Pass	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail
Number of Passes	1	12	6	6	6	1	7	9	0	1	3	6
% of CVD who Pass	3.44	41.38	20.69	20.69	20.69	3.44	24.14	31.03	0.00	3.44	10.34	20.69

Color vision deficient (CVD) subjects 1, 3, 8, 9, 11, 16, 17, 19, 21, and 26 failed all tests.

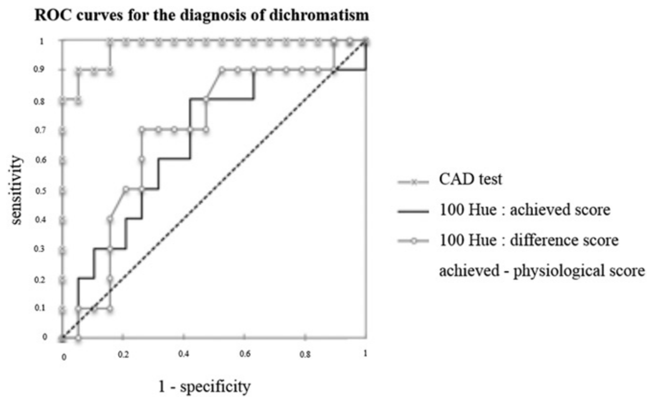

Fig. 1. ROC curves of the CAD and 100 Hue tests for the diagnosis of dichromatism.

DISCUSSION

The CVD population of our study consisted of 62% anomalous trichromats (37.9% deuteranomalous and 24.1% protanomalous trichromats) and 34% dichromats (20.7% protanopes and 17.3% deuteranopes). We had a higher proportion of protans and dichromats than in the average color-deficient population.[3]

The Farnsworth D15 test is currently the reference for Canadian Army pilots and Canadian civil aviation, as an alternative to lantern tests.[12] In our study, the use of this test would allow nine (45.5%) deuteranomalous trichromats, five (71.4%) protanomalous trichromats, and two (33.3%) protanopes to receive a class 1 medical certificate and, therefore, does not seem to be a satisfactory test. Among the various arrangement tests, the Lanthony 15 Hue test appears most relevant for color vision assessment as it detected all dichromats. In our study, its use would have led to the delivery of class 1 medical certificates to 36.4% of deuteranomalous trichromats and 28.6% of protanomalous trichromats. It is easy to implement and interpret.

The results of the different lantern tests are also inconsistent. For the Beyne lantern test, the French civil aviation protocol would allow a class 1 medical certificate to be delivered to 31% of color-deficient subjects, including one protanope and one deuteranope.

The 1 s / 4' protocol of the Beyne lantern test could be a possible alternative. The success rate for obtaining class 1 medical certificates is 20% and no dichromats passed this protocol. The $1/25^{th}$ s / 2' protocol is more selective and almost half of the CVN subjects failed. The most common error was the nonvisualization of the blue light. This can be explained by the physiological foveolar tritanopia.[23]

The Fletcher lantern test seems more effective for the detection of color-deficiency; however, it does not allow for differentiating minor impairments from more severe deficiencies. The difficulty of this lantern test seems to be due to:

- The simultaneous presentation of two colored lights decreases the capacity to distinguish between colors;
- The size of the light stimuli is much smaller than in the Beyne lantern; and
- The colors are difficult to recognize, with a yellow-green and a bright red that are close to white.

In 2016, Walsh evaluated color vision tests on 65 CVD subjects and 68 CVN subjects in the U.S. Army.[21] Using the anomaloscope as a gold standard, the authors reported a sensitivity of 0.86 and a specificity ranging from 0.85 to 1.00 for the CAD test. The Farnsworth D15 test showed a sensitivity of 0.35 and specificity of 1.00. Our results are consistent with Walsh's study except for a better diagnostic efficiency with the CAD test. The author concluded that the computerized cone contrast test (CCT) and CAD test performed well in terms of efficacy, with the CCT showing better sensitivity and specificity.

In a British study published by Squire et al. in 2005, the authors compared three lantern tests (Beyne, Spectrolux, and Holmes-Wright type A) and the Nagel anomaloscope in 55 CVD subjects and 24 CVN subjects.[20] In their study, the Ishihara plates showed a 1.00 sensitivity and 0.71 specificity. Unlike our study, where almost all CVN subjects succeeded all tests, the study by Squire et al. reported that 12 of 24 healthy subjects failed the Beyne lantern test and Nagel anomaloscope. Regarding the anomaloscope results, the authors considered matching ranges greater than 4 units as fails, even for normal trichromats. Of 12 normal trichromats who failed, 11 had a matching range of 5 units. In our study, we considered all subjects with a diagnosis of "normal color vision" using the anomaloscope as suitable, regardless of the width of matching range. Regarding the Beyne lantern, 92% of normal trichromats failed the test because they described the white light as "yellow." In our study, subjects were informed that white was an "off-white," which limited mistakes. As in our study, subjects that pass one secondary test are not guaranteed to pass the other tests. The authors concluded that the results for the tests authorized in the EASA standards present a high level of variability and inconsistency.

The most important study assessing the CAD test and lantern tests is the British CAA report, produced in collaboration with the Federal Aviation Administration (FAA): "Minimum Colour Vision Requirements for Professional Flight Crew; Recommendations for New Colour Standard Vision."[4] The authors evaluated 117 CVD subjects for their ability to recognize PAPI lights and their CAD test results. Of 77 deuteranomalous trichromats, 34 succeeded in the PAPI simulator. Among them, 29 had a CAD test score $<$6 SN. Out of 40 protanomalous trichromats, 20 subjects passed the PAPI test, including 13 with a CAD test score $<$12 SN. None of the subjects with a CAD test score lower than the proposed standards failed the PAPI test. In the CAA study, 36.1% of 255 deuteranomalous trichromats and 29.8% of 131 protanomalous trichromats passed the CAD test. In our study, the success rate for obtaining the class 1 medical certificate following the CAD test were lower. The difference in our results could be explained by a lower proportion of light anomalous trichromats. As in our study, the CAA study reported a $>$99% correlation between the CAD test and the anomaloscope for qualitative diagnosis.

Ryan Brookes from the New Zealand Defense Force also published a report in 2015 on color vision requirements.[7]

According to the author, the currently used clinical diagnostic tests (Farnsworth D15 and Nagel anomaloscope) are not suitable for professional aircraft crew. The author estimated that color-vision deficiency would not affect the ability of an applicant to operate an aircraft: the information coding is redundant and color signals in cockpits and PAPI are chosen to be recognized by color-deficient subjects.

In 2014, the FAA published a study that also evaluated the recognition of PAPI signals, comparing the use of light-emitting diodes (LEDs) with the currently used incandescent lamps.[18] Color-deficient subjects recognized the signals of incandescence lamps in the PAPI simulator as well as normal subjects and achieve even better results with LED lamps. Only subjects with red-green and yellow-blue deficiencies performed less well with the PAPI simulator. However, the severity of the color deficiency was not evaluated in this study.

In 2008, Cole compared the results of the Farnsworth lantern test and naming PAPI lights test in 52 CVD subjects and 52 CVN subjects.[8] A total of 10 color-deficient subjects who had passed the Farnsworth lantern test without errors made a great many more mistakes for the PAPI test than control subjects, and 80% made more errors than the worst-performing CVN subject. This study challenged the capacity of subjects who passed the lantern test.

The Fletcher lantern, as the Holmes-Wright lantern, presents two vertically stacked lights with a 0.9-min arc aperture. The colors used differ slightly, however, on the CIE diagram. In 2005, Fletcher assessed the Fletcher-Evans CAM lantern test; the results are similar to those of our study: 9 of 71 healthy subjects made mistakes during the first run and all the 18 color-deficient subjects failed at the Fletcher lantern test.[10]

The Fletcher-Evans CAM lantern is very sensitive for the detection of color deficiencies. It is therefore a good clinical test, but its relevance as a test to select for color vision is being discussed. It should not be compared with the Holmes-Wright lantern test, which leads to the delivery of class 1 medical certificates for 9–30% of color-deficient candidates, depending on studies.[4,14,20]

An American study published by Gaska and Wright in 2016 evaluated the recognition of colors in the cockpit for 45 CVN pilots and 49 color-deficient pilots.[11] They compared CCT results with color recognition in a "Situation Awareness" simulator and evaluated the accuracy, speed, and throughput of answers. The color-deficient subjects performed statistically less well than control subjects and a statistically significant relationship was reported between their performance and the CCT score.

Candidates try to minimize their color-vision deficiency. Some memorize the Ishihara plates and, even when presented out of order, can sometimes recognize them (depending on the layout, size, or saturation). The testing protocol should be adhered to strictly to ensure reliable and reproducible assessment, and the explanations given to the candidate must be protocolized. The anomaloscope, although easy to operate in automatic mode, in manual mode requires a good knowledge of the device itself and of color vision physiology.

Testing time is an important parameter in a screening context. The Ishihara test and the lantern tests are the simplest and fastest to implement. Both D15 tests take between 2 to 3 min, but the 100 Hue test is much longer (15 min). The anomaloscope is often performed in both automatic and manual mode, totaling about 20 min. The CAD test is fast (<5 min) for CVN subjects if the "fast screening" step is successful. In our study, six CVN (26%) had to continue on to the "Definitive CAD" step, which lasts about 12 min. In 2009, the CAA report recommended the use of the CAD test alone, without the Ishihara test. In their study, only 5% of healthy subjects had to continue to the "Definitive CAD" step.[4]

Some tests may be examiner-dependent. For the Fletcher-Evans CAM lantern test, the opening time is determined manually and so is not entirely reproducible. The anomaloscope in manual mode is also examiner-dependent.

This is one of the advantages of the CAD test, which is reproducible regardless of the examiner. In addition, the CAD test cannot be memorized and, therefore, can be repeated. Another advantage of the CAD test is to assess a greater number of colors: 16 chromaticities are tested, covering the red-green and blue-yellow deficiencies. The main disadvantage of the CAD and CCT tests is their higher price.

The moderate number of subjects in our study may limit its scope and could explain certain differences compared with previous studies. Four different examiners conducted the tests, but all were trained. The explanations and tests conditions were similar regardless of the examiner because all test protocols were standardized. The proportion of men and women was different in the two groups, but we do not believe that this affected the study results. The mean age of the CVN group was higher and could have influenced the results, but the difference is not statistically significant.

Our study evaluated ergonomic and clinical color vision tests. We did not use aviation light simulators. An ideal test would consist in recognizing lights in real flight conditions. The colors used in cockpits were also little assessed, as in previous studies. It would be interesting to compare the results of color vision tests with the recognition of colored elements in real cockpits, in different lighting conditions, in order to assess the impact of color-deficiency (decreased reaction time in case of danger, poor map reading).

In conclusion, the results, use, and simplicity of interpretation of the Ishihara test confirm its interest as a screening test for detecting red-green axis color-deficiency.[6] The Fletcher-Evans CAM lantern protocol is very restrictive. The Beyne lantern test has the advantage of being simple and fast to use, but some dichromats can recognize the colored lights without error. The Nagel anomaloscope provides an accurate measurement of color-deficiency and separates dichromats from anomalous trichromats. However, it is a lengthy test and requires highly experienced examiners, which could be a limit to its use.

The CAD test has the advantage, like the anomaloscope, of quantifying color-deficiency and using primary light sources. The pass criteria used by the British Civil Aviation Authority for pilots seems more stringent than the anomaloscope. CAD test results are encouraging, but its duration and cost could be a limitation for widespread use in the medical assessment of aircraft crew.

The multiplication of colored signals in new-generation cockpits and the use of a greater number of different colors may call for more stringent color vision requirements for pilots. The discrepancy in results confirm that the current color vision test protocols need to be reassessed. Furthermore, the current protocol does not detect yellow-blue deficiency. In terms of color vision selection of professional flight crew, the acceptable levels of color-deficiency still need to be defined at both the European and worldwide level to ensure flight safety.

ACKNOWLEDGMENTS

The authors have no conflicts of interest.

Authors and affiliations: Marie Marechal, M.D., Maxime Delbarre, M.D., and Françoise Froussart-Maille, M.D., Ph.D., Department of Ophthalmology, and Joaquim Tesson, Carole Lacambre, and Hélène Lefebvre, Aeromedical Center, Percy Military Hospital, Clamart, France.

REFERENCES

1. AMC and GM on the medical certification of pilots and medical fitness of cabin crew.pdf [Internet]. [Accessed 2017 July 30]. Available from: https://www.easa.europa.eu/system/files/dfu/AMC%20and%20GM%20on%20the%20medical%20certification%20of%20pilots%20and%20medical%20fitness%20of%20cabin%20crew.pdf.
2. Arrêté du 27 janvier 2005 relatif à l'aptitude physique et mentale du personnel navigant technique professionnel de l'aéronautique civile (FCL 3) [Order of 27 January 2005 on the physical and mental fitness of professional flying personnel in civil aeronautics]. JORF no. 61; 2005:4354. Paris (France): Minister of Equipment, Transport, Spatial Planning, Tourism and the Sea, and the Minister of Defense; 2005.
3. Barbur JL, Rodriguez-Carmona M. Colour vision requirements in visually demanding occupations. Br Med Bull. 2017; 122(1):51–77.
4. Barbur JL, Rodriguez-Carmona M, Evans S, Milburn N. Minimum colour vision requirements for professionnal flight crew. Recommendations for new colour vision standards. Safety Regulation Group. Civil Aviation Authority; 2009. 2013. [Accessed 2017 July 30]. Available from www.caa.co.uk/Aeromedical-Examiners/Medical-standards/Pilots-(EASA)/Conditions/Visual/Colour-vision-guidance-material-GM.
5. Birch J. Congenital protan and deutan defects in women. In: Drum B, Verriest G, editors. Colour Vision Deficiencies IX, vol. 52. Dordrecht: Doc. Ophthalmol. Proc. Series; 1989.
6. Birch J. Efficiency of the Ishihara plates for identifying red-green color deficiency. Ophthalmic Physiol Opt. 1997; 17(5):403–408.
7. Brookes RE. Colour vision requirements for aircrew [Internet]. New Zealand Defence Force. Defence Technology Agency; 2015. [Acessed 2017 Aug. 16]. Available from: http://www.dta.mil.nz/wp-content/uploads/Report405.pdf.
8. Cole BL, Maddocks JD. Color vision testing by Farnsworth lantern and ability to identify approach-path signal colors. Aviat Space Environ Med. 2008; 79(6):585–590.
9. Cruz-Coke R. Color blindness. An evolutionary approach. Springfield: Thomas; 1970.
10. Fletcher R. The Fletcher CAM lantern colour vision test. Optom Today. 2005; 29:24–26.
11. Gaska JP, Wright ST, Winterbottom MD, Hadley SC. Color vision and performance on color-coded cockpit displays. Aerosp Med Hum Perform. 2016; 87(11):921–927.
12. Government of Canada Nationale Defence. Instruction for Testing Colour Vision. Medical standards. Canadian Armed Forces [Internet]; 2013. [Accessed 2017 July 30]. Available from: http://www.forces.gc.ca/en/about-policies-standards-medical-occupations/cf-colour-vision-testing-instructions.page.
13. Hasrod N, Rubin A. Defects of colour vision: a review of congenital and acquired colour vision deficiencies. African Vision and Eye Health. 2016; 75(1):a365.
14. Hovis JK. Repeatability of the Holmes-Wright type A lantern color vision test. Aviat Space Environ Med. 2008; 79(11):1028–1033.
15. Lanthony P. Evaluation of the desaturated Panel D-15. I. Method of quantification and normal scores. J Fr Ophtalmol. 1986; 9(12):843–847.
16. Maille M, Crepy P. Les nouvelles normes visuelles en aéronautique militaire [New visual standards in military aeronautics]. Medecine Aeronautique Spatiale. 1990; XXIX:261–262.
17. Menu J-P. North Atlantic Treaty Organization, Research and Technology Organization, Human Factors and Medicine Panel. Operational colour vision in the modern aviation environment. France: RTO/NATO; 2001; [Internet]. [Accessed 2017 Aug. 24]. Available from https://www.sto.nato.int/publications/STO%20Technical%20Reports/RTO-TR-016/TR-016-$$ALL.pdf.
18. Milburn N, Gildea K, Perry D, Roberts C, Peterson L. Usability of light-emitting diodes in precision approach path indicator systems by individuals with marginal color vision [Internet]. Civil Aerospace Medical Institute, Federal Aviation Administration; 2014. [Accessed 2017 July 30]. Available from: https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2010s/media/201406.pdf.
19. Rabin J, Gooch J, Ivan D. Rapid quantification of color vision: the cone contrast test. Invest Ophthalmol Vis Sci. 2011; 52(2):816–820.
20. 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(5):421–429.
21. Walsh DV, Robinson J, Jurek GM, Capó-Aponte JE, Riggs DW, Temme LA. A performance comparison of color vision tests for military screening. Aerosp Med Hum Perform. 2016; 87(4):382–387.
22. Watson DB. Lack of international uniformity in assessing color vision deficiency in professional pilots. Aviat Space Environ Med. 2014; 85(2):148–159.
23. Williams DR, MacLeod DI, Hayhoe MM. Foveal tritanopia. Vision Res. 1981; 21(9):1341–1356.