Federal Aviation Administration

DOT/FAA/AM-09/11
Office of Aerospace Medicine
Washington, DC 20591

Minimum Color Vision Requirements for Professional Flight Crew, Part III: Recommendations for New Color Vision Standards

John Barbur
Marisa Rodriguez-Carmona
Applied Vision Research Centre
Northampton Square
London, EC1V 0HB, UK

Sally Evans
Civil Aviation Authority
Gatwick Airport South
Gatwick, RH6 0YR, UK

Nelda Milburn
Civil Aerospace Medical Institute
Federal Aviation Administration
Oklahoma City, OK 73125

June 2009

Final Report

OK-09-0434-JAH

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-09/11		

4. Title and Subtitle Minimum Color Vision Requirements for Professional Flight Crew, Part III: Recommendations for New Color Vision Standards	5. Report Date June 2009 6. Performing Organization Code
7. Author(s) Barbur J[1], Rodriguez-Carmona M[1], Evans S[2], Milburn N[3]	8. Performing Organization Report No.
9. Performing Organization Name and Address [1]Applied Vision Research Centre, Northampton Square, London EC1V 0HB, UK [2]Civil Aviation Authority, Gatwick Airport South, Gatwick, RH6 0YR, UK [3]FAA Civil Aerospace Medical Institute, P.O. Box 25082, Oklahoma City, OK 73125	10. Work Unit No. (TRAIS) 11. Contract or Grant No. 06-G-003
12. Sponsoring Agency name and Address Office of Aerospace Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591	13. Type of Report and Period Covered 14. Sponsoring Agency Code

15. Supplemental Notes

Work was accomplished under approved task AM-521.

16. Abstract

This report describes the findings of the third phase of the project sponsored by the United Kingdom Civil Aviation Authority on "Minimum Color Vision Requirements for Professional Flight Crew." This third part of the project, "Recommendations for New Color Vision Standards," involved collaboration and co-sponsorship by the Federal Aviation Administration.

Minimum color vision requirements for professional flight crew have been established by assessing the level of color vision loss above which subjects with color deficiency no longer perform the most safety-critical, color-related tasks within the aviation environment with the same accuracy as normal trichromats.

The new CAD (Color Assessment & Diagnosis) test provides accurate assessment of the applicant's color vision. The results of the test establish with high specificity whether the subject's red-green and yellow-blue color vision performance falls within the normal range and the class and severity of color vision loss in subjects with color deficiency. The results of the test also indicate whether the applicant's color vision meets the minimum requirements for safe performance that have emerged as necessary from this investigation.

If the new, experiment-based, pass/fail color limits were adopted as minimum requirements for professional flight crew, 36% of deutan subjects and 30% of protan subjects would be classified as safe to fly. Given the higher prevalence of deutan deficiencies, these findings suggest that 35% of color deficient applicants would be classified as safe to fly.

17. Key Words	18. Distribution Statement
Color Vision, Pilot Screening, Colour Deficiency, CAD Test, PAPI Lights, Signal Lights	Document is available to the public through the Defense Technical Information Center, Ft. Belvoir, VA 22060; and the National Technical Information Service, Springfield, VA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	47	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

ACKNOWLEDGMENTS

We thank Dr. Anthony Evans (ICAO) for his active participation in this project and continued support with many aspects of this investigation. We also acknowledge the valuable contribution each of the persons listed below has made to the project through progress review meetings, valuable criticism, and comments on the final report:

Ms. J. Birch	City University, UK
Mr. A. Chorley	CAA Chief Optometrist, UK
Dr. D. Connolly	Qinetiq Ltd., UK
Dr. H. Courteney	CAA, UK
Dr. A. Evans	ICAO Montreal, Canada
Mr. S. Griffin	CAA, UK
Mr. A. Harlow	City University, UK
Dr. I. Moorhead	Qinetiq Ltd., UK
Ms. M. O'Neill-Biba	City University, UK
Dr. J. Pitts	CAA Consultant Ophthalmologist, UK
Dr. R. Scott	RAF Consultant Ophthalmologist, UK
Mr. S. Williams	CAA, UK

EXECUTIVE SUMMARY

This report describes the findings of the third phase of the project sponsored by the United Kingdom Civil Aviation Authority (CAA) on "Minimum Color Vision Requirements for Professional Flight Crew." Parts I and II have already been published and cover "Use of Color Signals and Assessment of Color Vision Requirements," and "Visual Task Analysis," respectively. This third part of the project, "Recommendations for New Color Vision Standards," involved collaboration and co-sponsorship by the Federal Aviation Administration (FAA).

Minimum color vision requirements for professional flight crew have been established by assessing the level of color vision loss above which subjects with color deficiency no longer perform the most safety-critical, color-related tasks within the aviation environment with the same accuracy as normal trichromats. The new CAD (Color Assessment & Diagnosis) test provides accurate assessment of the applicant's color vision. The results of the test establish with high specificity whether the subject's red-green and yellow-blue color vision performance falls within the normal range and the class and severity of color vision loss in subjects with color deficiency. The results of the test also indicate whether the applicant's color vision meets the minimum requirements for safe performance that have emerged as necessary from this investigation. If the new, experiment-based, pass/fail color limits were adopted as minimum requirements for professional flight crew, 36% of deutan subjects and 30% of protan subjects would be classified as safe to fly. Given the higher prevalence of deutan deficiencies, these findings suggest that 35% of color deficient applicants would be classified as safe to fly.

BACKGROUND

The use of color in aviation for coding of signals and information is important, hence the need to set adequate color vision requirements to ensure that flight crew are able to discriminate and recognize different colors, both on the flight deck and externally. Concern has, however, been expressed during the past few years that the current color vision standards, at least within JAA (Joint Aviation Authorities) member states, are not appropriate since most tests and pass limits only screen for normal red/green color vision. Since the incidence of congenital yellow/blue deficiency is extremely low (see Table 1), the absence of red/green deficiency is virtually equivalent to normal trichromacy. Subjects with minimal color deficiencies often fail normal trichromacy tests, and the great majority are therefore prevented from becoming pilots, although many of these subjects may well be able to perform safety-critical tasks, as well as normal trichromats, when presented with the same, suprathreshold color signals. In principle, these subjects should be allowed to fly. To include some individuals with minimum color deficiency that may well be safe to fly, some authorities have either relaxed the pass limits on tests designed to screen for normal color vision (e.g., Ishihara, Dvorine) or they have introduced less demanding tests that applicants with mild color vision deficiency can pass. This approach does justice to some applicants but not to others. Existing, conventional color screening tests employed by most authorities cannot be used to quantify accurately the severity of color vision loss, and this makes it difficult to set reliable pass/fail limits. With very few exceptions, no red/green color deficient applicants pass either the Ishihara or the Dvorine color screening tests with zero errors. The same applies to anomaloscope matches when strict criteria are employed (e.g., when the applicant sets an appropriate red/green mixture field to match the color appearance of a yellow, monochromatic field, as in the Nagel anomaloscope). In this respect, these tests are excellent, but as has been shown in several studies, neither the anomaloscope results (Barbur et al., 2008) nor the Dvorine/Ishihara plates (Squire et al., 2005) can be used to quantify reliably the severity of color vision loss.

When the pass limits are relaxed, the outcome of such tests no longer guarantees normal trichromatic performance in the most safety-critical, color-related tasks. The FAA guidelines for aviation medical examiners (www.faa.gov/about/office_org/headquarters_offices/avs/offices/aam/ame/guide/) in relation to color vision tests, testing procedures, and pass limits are different from those practiced in Europe. The JAA member states employ the Ishihara screening test to identify applicants with red/green deficiency. No errors on any of the first 15 plates of the 24-plate Ishihara set are allowed in order to pass. Most color deficient observers (both deutan and protan) fail this stringent use of the Ishihara test, except for a very small number of minimum deuteranomalous that pass. In addition, ~15% of normal trichromats (an estimate based on 202 normal trichromats examined at Applied Vision Research Centre (AVRC)) also fail the Ishihara, when one employs the strict CAA/JAA pass/fail criteria. Secondary tests such as the Holmes-Wright Type A (HW) lantern (used in the United Kingdom) are employed, and although some color deficient observers (mostly deuteranomalous subjects, see discussion section) pass these tests, the severity of their color vision loss remains unknown. One advantage of using the HW lantern as a secondary test is that normal trichromats pass this test and are therefore not disadvantaged when they fail the Ishihara test. Mild deuteranomalous subjects that pass the lantern tests cannot be distinguished from normal trichromats on the basis of these tests. These subjects (i.e., the deutan applicants that pass the HW lantern) are therefore allowed to fly. All protan subjects fail the HW test, but some may have sufficient chromatic sensitivity to carry out safety-critical color tasks, as well as normal trichromats. It is, therefore, clear that, at least in the UK (which employs the HW lantern as the secondary test), protans are very likely to be excluded.

The current procedures within JAA are therefore unsatisfactory for at least two reasons. First, there is no guarantee that the deutan subjects that pass secondary tests can cope with safety-critical, color-related tasks since the severity of their color vision loss remains unquantified. Second, many color deficient subjects that can carry out such tasks safely fail the lantern tests and will not be allowed to fly. There are also other additional problems. The pass/fail variability of different conventional, color screening tests is high (Squire et al., 2005). Although subjects with minimum color deficiency may sometimes pass these tests, the results provide no reliable information as to the minimum color vision requirements that can be considered safe within the aviation environment.

Another important, practical aspect of regulatory testing of color vision is that aspiring pilots are often highly motivated to pass a screening test. The context in which the test is undertaken is therefore very different to the clinical setting. It is known that, in order to pass the Ishihara test, or similar pseudoisochromatic tests, color deficient applicants have been known to have learned the correct responses, so as to maximize their chances of passing the test. It is, for that reason, important to eliminate any opportunity of learning the right answers to pass a screening test. When used in the recommended clinical settings, most of the popular occupational color tests exhibit large within-subject and inter-subject variability, even within normal trichromats (Squire et al., 2005). The recommended surround, ambient viewing conditions, measurement procedures, and interpretation of results can vary significantly from country to country, even when the same tests are employed. Many ICAO member states have different requirements for color vision assessment and use different tests. Within the JAA, the HW (United Kingdom), the Spectrolux (Switzerland), the Beyne (France, Belgium, and Spain) lanterns, and the Nagel anomaloscope (Germany) are recognized secondary tests. Since the correlation between the outcomes of different tests is poor (Squire et al., 2005), it is not uncommon for pilot applicants to fail the color vision assessment in one country and to pass in another. Although such occurrences have not passed unnoticed, no adequate solutions have emerged to set minimum limits of color vision sensitivity that can be considered operationally "safe" within specified environments. The lack of adequate solutions to this problem explains why some authorities unknowingly insist on normal trichromacy (which is largely what is achieved when current pass/fail limits are employed with Ishihara as the primary test and HW lantern as the secondary test).

The FAA guidelines are more liberal and allow for the use of various pseudoisochromatic plates (e.g., Ishihara, Dvorine, AOC-HRR, Richmond 1983-edition, 15 plates, etc.) with relaxed pass/fail limits (e.g., the applicant has to make seven or more errors in order to fail the Ishihara or Dvorine tests). Alternative tests such as Farnsworth lantern (FALANT), Keystone Orthoscope, etc., can also be used as acceptable substitutes. As far as the FAA guidelines are concerned, the analysis of results is restricted to the FAA approved tests that have been included in this investigation (i.e., Ishihara and Dvorine). In addition, the Aviation Lights Test (i.e., a modified Farnsworth lantern that was developed to screen air traffic controllers, see Fig. 7 in full report) was also included in this investigation.

NEW DEVELOPMENTS

Advances in understanding human color vision (Barbur, 2003) and the development of novel methods to measure accurately the loss of chromatic sensitivity (Barbur et al., 1994) have prompted the CAA to sponsor new studies to examine how color vision loss can be measured accurately and also to establish minimum color vision requirements for civil aviation professional pilots. As a result of the progress made in these studies, it is now possible to define the variability that exists within normal color vision and to detect with confidence and classify accurately even the smallest congenital color vision deficiencies that sometimes pass undetected in conventional, occupational color vision tests. More importantly, it is now possible to achieve the aim of the project, i.e., *to quantify the severity of color vision loss and to recommend minimum color vision requirements by establishing the level of color vision loss when color deficient observers can no longer perform the most safety-critical, color-related tasks with the same accuracy as normal trichromats.*

A number of developments that have emerged from the studies carried out during the last few years have made it possible to achieve the aim of this project:

- A Color Assessment and Diagnosis (CAD) test that employs novel techniques to isolate the use of color signals and measures accurately both red-green (RG) and yellow-blue (YB) chromatic sensitivity has been developed and validated (see Fig. 9).

- A study that compared outcome measures in the most common, occupational color vision tests, in both normal trichromats and in a large number of color deficient observers, has improved our understanding of current limitations. The findings from this study also justify the need for a test that can be used to accurately measure the subject's chromatic sensitivity and the variability expected within the color normal population (Squire et al., 2005).
- The establishment of color discrimination limits for normal vision (i.e., the standard normal CAD observer based on RG and YB color detection thresholds measured in ~250 normal trichromats) provides a template for detection of abnormal sensitivity (Rodriguez-Carmona et al., 2005; see Fig. 8). In addition, similar measurements in over 300 color deficient observers that participated in several projects related to color vision provided the data needed to describe the differences in the severity of color vision loss within deuteranomalous and protanomalous observers (see Fig. 12).
- Identification of the most important, safety-critical, color-related tasks for pilots, and faithful reproduction of such tasks in the laboratory made it possible to establish experimentally the safe limits of color vision loss. The visual task analysis carried out as part of this study identified the Precision Approach Path Indicator (PAPI) as the most important, safety-critical task that relies largely on color vision. At some airports, color signals are also used for aligning the aircraft when approaching the parking area, and in such cases, correct color recognition is critical to the safe accomplishment of this task. There are many other tasks that involve the use of color signals, but they involve larger stimuli. The viewing is under more favorable conditions of light adaptation and other cues make the color coding less critical. In the case of the PAPI lights, it is essential that the pilot distinguishes accurately the number of "white" and "red" lights. Moreover, the pilot needs to recognize the four adjacent lights as "white," when too high, and as "red," when too low. The PAPI lights task is demanding since the lights can be very small (i.e., subtend a very small visual angle at the eye) and are often seen against a dark background (see Fig. 18) when color discrimination sensitivity is known to be poor.
- Color discrimination limits (based on the CAD test) that can be classified as safe for pilots in civil aviation have been established. This was achieved by measuring and relating PAPI task performance and color discrimination sensitivity as assessed on CAD, signal lights, and a number of other color vision tests (see description below) in 40 protanomalous, 77 deuteranomalous, and 65 normal trichromats. There are other visual tasks that can be classed as safety-critical, but these generally involve larger and brighter lights, and are therefore easier to carry out. These tasks either rely on color discrimination (such as the red-green parking lights) or, in some cases, the tasks benefit from the use of color signals as redundant information (such as the "green" runway threshold lights). In addition to the red-green parking and the green runway threshold lights, there are also a number of other runway lights: the red-white centerline lights, the green-yellow lead-off lights, and the red stopway lights. The tasks that involve these additional lights have not been simulated in the laboratory, but, as argued in the main report, they are either less demanding in terms of color discrimination, or the color signals are only used to reinforce the functional significance of the lights.
- Data showing correlation between PAPI scores and CAD sensitivity thresholds are shown in Fig. 23 for normal, deuteranomalous, and protanomalous observers.

PRINCIPAL CONCLUSIONS

- Subjects with red/green congenital color deficiency exhibit an almost continuous loss of chromatic sensitivity. The loss of sensitivity (when expressed in Standard Normal (CAD) units (SN units) is greater in protanomalous than deuteranomalous observers (Fig. 12).
- When the ambient level of light adaptation is adequate, normal aging does not affect significantly either RG or YB thresholds below 60 yrs of age (see Fig. 13).
- Analysis of PAPI results shows that the use of a modified "white" light results in significant, overall improvements in PAPI performance, particularly within normal trichromats and deuteranomalous observers. The modified (or color corrected white) is achieved simply by adding a color correction filter to the standard white lights produced by the source. The filter employed in this study decreased the color temperature of the standard white (used in PAPI systems) by 200 MIREDS (micro reciprocal degrees).
- The deuteranomalous subjects investigated in this study with CAD thresholds <6 SN units and the protanomalous subjects with CAD thresholds <12 SN units perform the PAPI test, as well as normal trichromats.
- 43 of the 77 deuteranomalous subjects failed the PAPI test. 29 out of the remaining 34 subjects that passed the PAPI test had CAD thresholds <6 SN units.
- 20 of the 40 protanomalous subjects failed the PAPI test. 13 out of the remaining 20 subjects that passed the PAPI test had CAD thresholds <12 SN units.

- A small number of deuteranomalous (5) and protanomalous (7) observers with thresholds higher than 6 and 12 SN units, respectively, passed the PAPI test. All these subjects do, however, exhibit poor overall RG chromatic sensitivity in all the other color tests employed in the study and are therefore likely to be disadvantaged in many other suprathreshold visual performance tasks that involve color discrimination.
- The results suggest that subjects with minimum color deficiency that does not exceed 6 SN units for deuteranomalous observers and 12 SN units for protanomalous observers perform the PAPI test as well as normal trichromats. If these findings were adopted as pass/fail limits for pilots ~35% of color deficient applicants would be classified as safe to fly.
- The administration of the CAD test eliminates the need to use any other primary or secondary tests. It is proposed that a rapid, reduced version of the CAD test (labelled fast-CAD) is administered first to establish whether the applicant passes with no errors the 6 SN limit established for deutan subjects. Deutans represent ~6% of color deficients and 36% of deutan subjects pass the recommended CAD limit (see Table 3). When one includes normal trichromats, ~94% of all applicants will pass the fast-CAD screening test and be classified as safe to fly. This process is very efficient since the fast-CAD test is simple to carry out and takes less than 30 seconds to complete. The definitive CAD test (which takes between 6 to 8 minutes for RG sensitivity) is administered only when the applicant fails the fast-CAD screening test. The latter establishes the class of color deficiency involved and whether the applicant's threshold is below the pass/fail limit established for protan subjects. In addition, the CAD test provides the option to test the applicants YB color vision. This option reveals whether the applicant's YB discrimination sensitivity falls within the normal range. In view of the increased use of color in aviation, testing for normal YB thresholds can also be of relevance to aviation safety.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

DEFINITIONS

For a relevant list of definitions refer to a previous CAA Paper 2006/04 (CAA 2006a).

Abbreviations

AGL -------- Aeronautical Ground Lighting
ALT--------- Aviation Light Test
ATC -------- Air Traffic Control
ATCS------- Air Traffic Control Specialist
AVRC ------ Applied Vision Research Centre (City University)
CAA -------- Civil Aviation Authority
CAD-------- Color Assessment and Diagnosis test
CIE --------- Commission Internationale de l'Eclairage
CS ---------- Chromatic Sensitivity
FAA--------- Federal Aviation Administration
ICAO ------ International Civil Aviation Organization
JAA --------- Joint Aviation Authorities
JAR --------- Joint Aviation Requirements
L-cones----- Long-wavelength sensitive cones
LC ---------- Luminance Contrast
MIRED ---- Micro Reciprocal Degrees
M-cones---- Medium-wavelength sensitive cones
NTSB ------ National Transportation Safety Board
PAPI -------- Precision Approach Path Indicator
RG ---------- Red-Green
S-cones ----- Short-wavelength sensitive cones
SI------------ Système International d'Unités (International System of Units)
SN ---------- Standard Normal
UK---------- United Kingdom
YB ---------- Yellow-Blue

Nomenclature

° - -- -- -- -- Degrees
cd m^{-2}- -- -- Candelas per square meter
λ- -- -- -- -- Wavelength (lambda), nm
λ_{max}- -- -- -- Maximum (peak) wavelength of $V(\lambda)$
% -- -- -- -- Percent
2' arc-- -- -- 2 minutes of arc
A- -- -- -- -- Ampere (amp) unit of electric current
km - -- -- -- Kilometer
K- -- -- -- -- Degrees Kelvin
mm- -- -- -- Millimeters (1 mm=10^{-3} of a metre)
M -- -- -- -- Mired (unit of measurement to express color temperature = 10^6 K)
nm - -- -- -- Nanometers (1 nm=10^{-9} of a metre)
s-- -- -- -- -- Second (time)
μ- -- -- -- -- Micro=x10^{-6}
$V(\lambda)$ -- -- -- Standard photopic luminous efficiency (for high ambient illumination) (CIE, 1924)
$V'(\lambda)$-- -- -- Standard scotopic luminous efficiency (when very low light levels are involved) (CIE, 1951)

Minimum Color Vision Requirements for Professional Flight Crew, Part III: Recommendations for New Color Vision Standards

1.0 INTRODUCTION

Occupational color vision standards were introduced in aviation in 1919 by The Aeronautical Commission of the International Civil Air Navigation Authority. These standards reflected both the needs and the methods available for color vision assessment at the time. Concern has been expressed during the last few years that the current JAR (Joint Aviation Requirements) color vision standards may be too stringent and, at the same time, also variable. The tests employed do not always reflect the tasks pilots encounter in today's aviation environment. An examination of current standards and techniques employed to assess color vision requirements suggests the need for a more unified color vision test to provide a measure of color vision loss that relates directly to the most safety-critical, color-related tasks within the aviation environment (Cole, 1993). The current color vision standards and accepted JAA (Joint Aviation Authorities) color vision tests for professional flight crew have been reviewed by the United Kingdom Civil Aviation Authority (UK CAA). This report follows other CAA documents published in 2006: "Minimum Color Vision Requirements for Professional Flight Crew: Part 1. The Use of Color Signals and the Assessment of Color Vision Requirements in Aviation, and Part 2. Task Analysis."

Aviation accidents have high social and economic costs, especially if the accident involves large passenger aircraft. Rigorous safety standards have been established over decades to decrease the probability of aviation accidents. An important strategy in achieving high levels of safety in aviation is to build redundancy in equipment and the interpretation of signals and other information by pilots and other personnel. Color is used extensively to code information in the aviation environment, and pilots are normally expected to have good color discrimination. Even when other cues are also available, the ability to use color information increases redundancy, and in some tasks, this improves considerably the level of visual performance that can be achieved. Some accidents have been linked to loss of color vision (National Transportation Safety Board, 2004). There is also some evidence to suggest that the likelihood of accidents is increased in pilots that are color deficient (Vingrys & Cole, 1986). Other studies have shown that subjects with color vision deficiencies make more errors and are slower in recognizing aviation signals and color coded instrument displays (Vingrys & Cole, 1986; Cole & Maddocks, 1995; Squire et al., 2005). There are also a small number of tasks in which color information is not used redundantly; therefore, the correct interpretation of color signals becomes very important.

1.1.1 The Use of Color in Aviation

The use of color in the aviation environment is important since it makes possible the efficient coding of signals and information and this, in turn, enhances visual performance, provided the subjects can make use of color signals. Humans with normal trichromatic color vision possess three distinct classes of cone photoreceptors. These contain short (S), middle (M), and long (L) wavelength sensitive photopigments with appropriate peak absorption wavelengths (λ_{max}). Variant L- and/or M-cone genes can cause significant shifts in λ_{max} and this, in turn can cause large changes in chromatic sensitivity. In addition to λ_{max} changes, other factors such as the amount of pigment present in photoreceptors can also affect chromatic sensitivity. Red/green deficiency is the most common type and is caused by either the absence of or the abnormal functioning of L- or M-cones. The corresponding condition is normally described as protan or deutan deficiency, respectively. Color vision deficiency affects approximately 8% of men and less than 1% of women (see table in section 4.4).

1.1.2 Current Color Vision Requirements and Assessment Methods in Aviation

The International Civil Aviation Organisation (ICAO) requires member nations to maintain a color vision standard to ensure pilots can recognize correctly the colors of signal lights used in aviation: "*The applicant shall be required to demonstrate the ability to perceive readily those colors the perception of which is necessary for the safe performance of duties*" (ICAO, 2001b). Many ICAO member states have different requirements for color vision assessment and employ different tests as the standard.

In Europe, there is agreement among the 38 members of the JAA to apply the same standards, at least in terms of primary tests. The current JAA color vision requirements (Section 1; JAR-FCL 3, 2002) use the Ishihara pseudo-isochromatic test (section 1.3.1 of this report) as a screening test for color vision. The JAA use the first 15 plates of the 24-plate version of the Ishihara pseudo-isochromatic test, with no errors as the pass criteria. If the applicant fails this test then either a lantern test or the Nagel anomaloscope test is used. The three lanterns recommended by the JAA are the Holmes-Wright Type A (United Kingdom), the Spectrolux (Switzerland), and the Beyne (France, Belgium and Spain). The subjects pass when they make no errors on the corresponding lantern test. For the Nagel anomaloscope (section 1.3.3 of this report), *"This test is considered passed if the color match is trichromatic and the matching range is 4 scale units or less…"* (See Appendix 14 to subpart B; JAR-FCL 3, 2000). The tests currently employed by JAA member states and the corresponding pass/fail criteria are fully described in the report by the CAA (Civil Aviation Authority, 2006a).

In the U.S.A., the Federal Aviation Administration (FAA) guidelines are more liberal, and the approved primary tests include Ishihara, Dvorine, AOC-HRR, Richmond, etc. The pass limits are also more relaxed which favors some applicants with mild color deficiency. Other tests include the Farnsworth Lantern, Keystone Orthoscope, etc. In exceptional cases, this can then be followed by the more practical Signal Light Gun Test (SLGT), usually carried out in an airport tower. This approach does justice to some applicants but not to others (see section 4). The disadvantage of this more liberal approach is that when the pass limits are relaxed, the pass/fail outcome of the screening tests no longer guarantees normal performance in the most safety-critical, color-related tasks.

Follow-up color vision tests may be carried out for renewal of Class 1 medical certificates. Within the JAA, the Ishihara plates which screen only for red/green deficiency are normally used for this assessment. Any yellow-blue loss (either congenital or acquired) will not therefore be picked up by this test (section 1.6 of this report). Since changes in chromatic sensitivity are often indicative of early-stage systemic (e.g., diabetes) or ocular diseases (e.g., glaucoma, age-related macular degeneration), it is recommended that both red-green and yellow-blue color sensitivity should be assessed with every medical examination and any significant changes noted. The data can then be used to detect when the progression of any inherent condition yields color thresholds that fall outside the range established for normal vision.

1.1.3 Problems Identified With Current Assessment Methods and Procedures

Current color vision requirements vary from country to country, even within the JAA member states. The correlation between the outcomes of different tests is poor; therefore, it is not uncommon for pilot applicants to fail the color vision assessment in one country and to pass in another (Squire et al., 2005). This is not completely unexpected given the large inter-subject variability, the different factors that can contribute to loss of chromatic sensitivity, and the different characteristics of the various color vision tests. The lack of standardization often causes confusion among applicants and provides the opportunity to attempt several tests in order to pass one of the many color vision standards.

1.1.4 A New Approach Based on Recent Advances in Color Vision Testing

Advances in the understanding of human color vision (Barbur, 2003) and the development of novel methods to measure accurately the loss of chromatic sensitivity (Barbur et al., 1994) have prompted the UK CAA and the FAA to sponsor new studies to examine how color vision loss can be measured accurately and to establish minimum color vision requirements for professional pilots. As part of the CAA funded study, the current accepted JAA color vision requirements for professional flight crew have been reviewed and the variability associated with the most common occupational color vision tests assessed, both in normal trichromats and in subjects with red/green color deficiency. The aim of the current project was to establish minimum limits of color vision sensitivity that can be considered to be operationally "safe" within the aviation environment. This joint report follows other CAA documents published in 2006 (CAA 2006/04): "Minimum Color Vision Requirements for Professional Flight Crew: Part 1. The Use of Color Signals and the Assessment of Color Vision Requirements in Aviation, and Part 2, Task Analysis."

1.1.5 A New Color Vision Test

Ideal color vision assessment requires a test that (i) provides true isolation of color signals and quantifies the severity of color vision loss, (ii) is based on data that describe the statistical limits of color discrimination in "*normal*" trichromats, so as to be able to differentiate minimal color vision loss due to congenital and acquired deficiencies from fluctuations expected within normal trichromats, (iii) has enough sensitivity to detect "minimal" deficiencies and to classify accurately the class of deficiency involved, and (iv) can be used to detect and monitor "significant changes" in color sensitivity over time. The Color Assessment and Diagnosis (CAD) test has been developed and improved over several years to fulfill these requirements (Section 1.4 of this report).

1.2 Identification of the Most Safety-Critical and Demanding Color Vision Tasks

An important aspect of this study was to investigate whether subjects with minimal color vision loss were able to carry out the most demanding, safety-critical, color-related tasks with the same accuracy as normal trichromats. If the findings indicate that "normal" color vision is not required to carry out such tasks, it then becomes important to establish the limits of color vision loss that can still be considered safe within the aviation environment.

The approach adopted in this investigation was to relate the accurate assessment of color vision loss to the subject's ability to carry out the safety-critical, color-based tasks within a specified environment when the use of cues other than color was minimized. A visual task analysis was carried out (CAA, 2006a) to identify and characterize the most important safety-critical, color-related tasks for flight crew. The PAPI and the parking signal lights were found to be the most safety-critical, color-related tasks when no redundant information is available to carry out the task. An earlier study by Cole and Maddocks (1995) also identified the PAPI lights as the most safety-critical, color-related task. The PAPI lights provided the pilot with accurate glide slope information on final approach to landing using four lights, each of which can be either red or white. Two whites and two reds indicate correct approach path, too many reds indicate that the approach height is too low, and too many whites indicate that the approach height is too high. The geometry of the PAPI signal system is shown in Fig. 1a (see also Fig. 2a).

An alternative system, the VASI (Visual Approach Slope Indicator), is sometimes used in North America and Australia. The VASI is more expensive and requires more space. There are several versions of the VASI system, but the main task of the pilot remains the discrimination of horizontal bars of well defined red and white lights (see Fig.1b). The more favorable geometry and the greater angular separation of the lights make the VASI color discrimination task less demanding than the PAPI. In the T-VASI version of the system, the changing geometry of the lights provides the required approach slope information, hence the color coding is used redundantly. The PAPI lights system is visually more demanding. The angular subtense of each of the four red/white lights corresponds to the smallest retinal image that can be produced by the

Figure 1a: The Precision Approach Path Indicator (PAPI) signal lights are used to inform pilots of the correct glide path for landing. This is an efficient system since it consists only of four lights, each of ~ 8" diameters, spaced 9m apart. The angular subtense of the lights remains relatively unchanged beyond ~ 800m, but the retinal illuminance of the lights decreases. The output of the lamps can be adjusted using 6 preset lamp current settings. The disadvantage is that the correlated color temperature of the lamp changes significantly and this affects mostly the color of the "white" lights (see section 2).

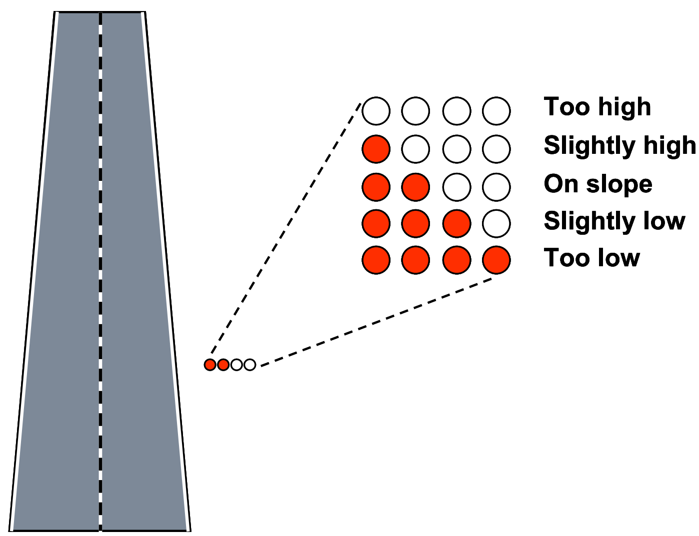

Figure 1b: The Visual Approach Slope Indicator (VASI) and the more sophisticated T-VASI systems. The correct angle of approach is indicated as red over white. The T-VASI requires more space and is more expensive, but the use of color provides only redundant information.

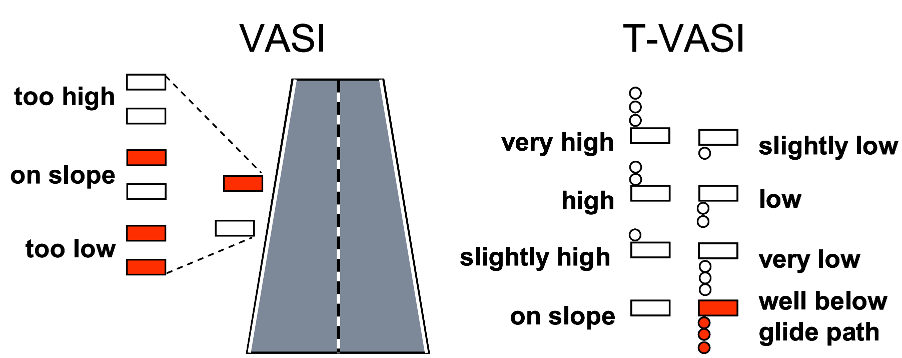

Figure 2a: Photograph of PAPI lights viewed from ground level. Photograph taken at Sussex Flight Centre, Shoreham Airport (December 2006).

optics of the eye, and the red/white color coding is used non-redundantly. This project has, therefore, focused on the PAPI lights as the most safety-critical, color-coded task for pilots.

1.2.1 Other Uses of Signal Lights Within the Aviation Environment

There are many other color signals that are used in the aviation environment to enhance conspicuity, code information, and group objects of interest together. These situations are less safety-critical, involve the use of larger stimuli under more favorable conditions of light adaptation, and the same information is also available in some other ways (e.g., text or audible signals). The PAPI signal system, on the other hand, offers no redundancy – at night there is no other unique cue in addition to color discrimination to help the pilot recognize the red and white light signals reliably to visually determine whether the aircraft is on the correct approach path for landing.

a. Parking Lights

Parking the aircraft requires the correct alignment of the aircraft with the line of approach. The pilot is aided in this task by the red-green parking lights. Both lights are seen as green when the aircraft is positioned correctly for approach. A red-green combination (see photograph of parking lights in Fig. 2b) signals that the aircraft has to veer slowly (towards the green light) to ensure that the pilot sees two green lights.

This is a color-related, safety-critical task simply because no other redundant cues are available; however, the lights are bright, the color difference between the lights is large, and the lights subtend a large visual angle at the eye. Consequently, the color discrimination task is less demanding, and it is expected that observers with minimum color vision deficiency may be able to carry out this task with the same accuracy as normal trichromats. This task was not investigated in this study.

b. Runway and Taxiway Lights

The lighting of runways and taxiways involves the use of color signals, but the use of color for coding is often redundant. The correct information is also provided by the geometry of the lights. Runway lighting is used for landing and take-off. On approach, the lighting of the runway provides essential information that yields outline views of the geometry of the runway. On touch-down, the lights form unique geometric lines and shapes that convey specific information. A particular runway may have some or all of the following lights:

- *Runway Edge Lights* are white (or amber) and run the length of the runway on each side.
- *Runway Threshold Lights* are green and indicate the starting point for the available landing distance.
- *Runway End Lights* are red and delineate the extremity of the runway that is available for maneuvering.
- *Runway Centerline Lights* start white, become red-white intermittent and then red only, towards the end of available runway for take-off.
- Touchdown Zone Lights consist of rows of white light bars (with three in each row) on either side of the centerline over the first 914 m of the runway (or to the midpoint, whichever is less).
- *Stopway Lights* are four unidirectional red lights equally spaced across the width to mark the end of any stopway associated with a runway used at night.

Runway edge lights provide perspective cues on approach and are less demanding than the PAPI light system; runway lighting becomes visible from several kms and often aids the pilot's visual search to locate the

Figure 2b: Photograph of the red and green parking lights that are used at airports to indicate to the pilot the correct line of approach for parking the aircraft at the stand. The pilot sees two green lights when the aircraft is positioned correctly for approach. A red-green combination (as shown in the photograph) signals that the aircraft has to veer slowly right (towards the green light) whereas a green on the left and a red on the right signals the need to veer slightly to the left (again towards the green light). This task description illustrates clearly that the pilot has to be able to discriminate between the red and green lights. The angular subtense of the parking lights is much larger than the PAPI and the retinal illuminance generated is also higher. This color discrimination task is therefore likely to be less demanding.

PAPI lights. The color of the runway threshold lights is largely redundant because these green lights cannot be confused with any other similar lights in terms of location, geometry, and shape, but the green color may reinforce their function. Touchdown zone lighting is added in order to improve texture and perspective and to give the flight crew an indication of the area within which a landing must be initiated. The geometry and location of the runway end lights in relation to other lights is, again, sufficient to indicate their function. The color of the runway centerline lights changes from "white" on touchdown to alternating "white-red" lights and then to "red" lights when the aircraft advances towards the end of the runway. The color of the lights indicates the position of the aircraft on the runway, and this information is important in some situations (e.g., when take-off has to be abandoned, especially in conditions of poor visibility). Runway signal lights tend to be larger and brighter than PAPI lights, and this makes the discrimination of color differences less demanding. It has therefore been assumed that if the applicants can discriminate the red and white PAPI lights from 5 kms, they should be able to easily discriminate the red and white lights on the runway centerline.

For night operations, taxiways at most airports are equipped with lights that may include some or all of the following:

- *Taxiway Edge Lighting* is blue to outline the edges of taxiways during periods of darkness or restricted visibility.
- *Taxiway Centerline Lighting* is green and provides centerline guidance on taxiways and aprons and when entering or vacating a runway.
- *Stop-bar Lights* are a single row of red, flush, or semi-flush inset lights installed laterally across the entire taxiway showing red towards the intended direction of approach. Following the controller's clearance to proceed, the stop-bar is turned off, and the centerline lead-on lights are turned on.
- *Runway Guard Lights* are either a pair of elevated flashing amber lights installed on either side of the taxiway, or a row of in-pavement yellow lights installed across the entire taxiway, at the runway holding position marking at taxiway/runway intersections.

Taxiway lights are seen from much shorter distances when the aircraft moves slowly on the ground. The discrimination by the pilot of the center taxiway line as green and the edge as blue is not an essential requirement to carry out the task safely, but the use of appropriate colors may well emphasise the function of the lights. Therefore, there is little doubt that an acceptable level of color discrimination is needed which can enhance the conspicuity of light signals, even when color is used redundantly and the tasks are less demanding or safety critical. The stop-bar and runway guard lights both play an important role in preventing runway incursions. In addition, the flashing aspect of the guard lights adds conspicuity to these signals but may also distract the pilot from interpreting other signals. The most common causes of runway incursions do not involve the incorrect interpretation of color signals since color is used redundantly and simply adds to the conspicuity of the lights. Other factors such as lack of communication between controller and pilot, lack of familiarity with airport layout, tiredness, lack of attention, and poor cockpit procedures for maintaining orientation in low visibility conditions (*ICAO NAM/CAR/SAM Runway Safety/Incursion Conference, Mexico City, October 2002*) can all contribute to runway incursions.

1.2.2 Analysis of the PAPI Lights Task

The PAPI task is a simple, efficient, red-white, two-color code (and the white and red lights generate both red-green and yellow-blue color signals in the eye). Red/green color deficient observers will continue to have full use of their yellow-blue channel, although the properties of this channel will differ between deutan and protan subjects. The PAPI system is efficient since it takes a small amount of space, and the size of the image of each light generated on the retina remains largely unchanged as the viewing distance increases beyond ~0.8 km (although the lights become less bright as the viewing distance is increased). The geometry of the lights carries no information and hence the need to use colored signals. It has been suggested that dichromats (who exhibit severe red/green color vision loss) may be able to correctly interpret differences between two colors, at least under some conditions (Heath & Schmidt, 1959). In addition, color deficient observers can usually recognize red signals with few errors (Vingrys & Cole, 1993; for a review, see Cole, 2004). It is likely that some subjects with severe color deficiency may be able to carry out the PAPI task with no errors, but it is essential to ensure that the subjects recognize and name all four lights as red when too low and as white when too high. Any simulation of the PAPI test must include all conditions and must also ensure that the use of brightness difference cues is minimized. On the other hand, the recognition of the red and white PAPI lights is not always an easy task. Atmospheric scatter and the use of reduced lamp current settings at night to dim the lights can shift the white signal toward the yellow region of the spectrum locus (see Fig. 17). This often causes problems for color normal observers and may cause even greater problems for color deficient observers. In the case of large passenger aircraft, the PAPI lights are seen from large distances (>5 kms) at night when both the angular subtense of each light and the angular separation between adjacent lights is very small. Adjacent lights tend to overlap visually, and this is particularly troublesome at night when the pupil size is large. Subjects with large, higher-order aberrations and increased light scatter in the eye will be disadvantaged at night. Although most subjects will have high visual acuity (<1 min arc) under photopic conditions, subjects with large higher order aberrations and scattered light may have very poor visual acuity under mesopic conditions when the pupil size is large. Visual acuity at low light levels in the mesopic range is not normally assessed for certification purposes. Partial overlap of adjacent lights makes the task of discriminating the number of red and white lights even more difficult. These additional factors explain why the PAPI task (which involves only two colors) is considered to be more critical than other color based tasks.

1.2.3 Disability Discrimination

There are further considerations that justify the need to establish safe, minimum requirements for color discrimination (when appropriate) and to avoid the easier alternative (from a regulatory viewpoint) of requiring every applicant to have normal color vision. The recent UK Disability Discrimination Act (2004) has, to a certain extent, exposed weaknesses in the current standards and procedures. Companies need to justify refusal to employ an applicant on the basis of his/her defective color vision, and this requires scientific evidence to demonstrate convincingly that the applicant will not be able to carry out necessary occupational tasks that involve color vision with the accuracy and efficiency expected of normal trichromats. In view of these arguments, we have developed a PAPI simulator and a PAPI Signal Lights test that can be used under controlled laboratory conditions. The simulators reproduce both the photometric and the angular subtense of the real lights under demanding viewing conditions when the lights are viewed against a dark background. The aim was to correlate the measured loss of chromatic sensitivity on the CAD test with the subject's performance on the most safety-critical, color-related task identified in the aviation environment. Since other color-related tasks (such as seeing the color of the parking lights or the discrimination of runway, centerline, and red and white lights) are less demanding, it is assumed that the pilot will also be able to correctly perform these tasks. In principle, this approach should make it possible to recommend pass/fail limits based on the observer's ability to carry out the most safety-critical and demanding PAPI task.

1.3 Brief Description of the Most Common Occupational Color Vision Tests

For a full description of color vision tests used in aviation, please refer to CAA Paper 2006/04 (2006a), and for a list of tests accepted by the FAA, see the FAA Guide for Aviation Medical Examiners (2008). The following color vision tests will be described here since they have been used along with the CAD test in this study. These are the Ishihara and Dvorine pseudoisochromatic plate tests, Nagel anomaloscope, and the Aviation Lights Test (ALT). Measures of color discrimination performance computed from the results of these tests will be examined and compared with the subject's scores on the PAPI simulator. The same PAPI simulator was also used to produce a more demanding signal lights test that required the subject to name one of six different colored lights, as described in section 2 of this report. The latter will be referred to as the PSL (PAPI Signal Lights) test.

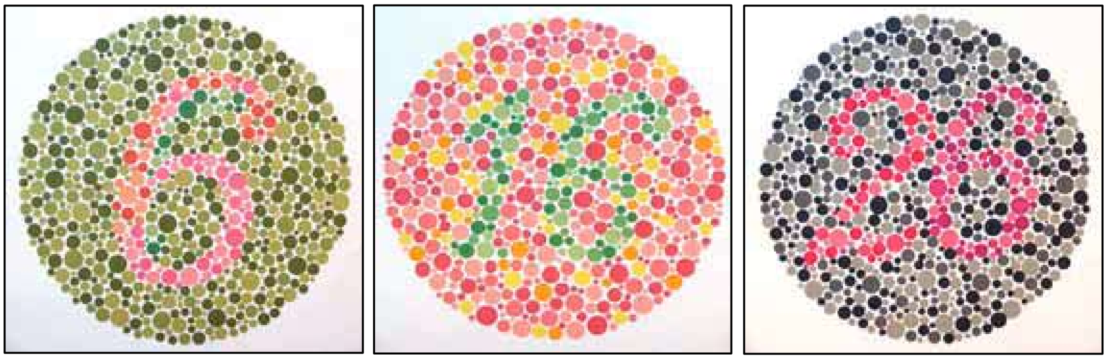

Figure 3: Ishihara pseudoisochromatic plates; left, transformation design; middle, vanishing design; right, protan/deutan classification plate. Please note that these may not be reproduced accurately as the printed color and the viewing illuminant will be different.

1.3.1 Ishihara Plate Test

The Ishihara pseudoisochromatic plate test consists of a series of numbers outlined by different colored dots as shown in Fig. 3. This is the most widely accepted screening test for red/green color deficiency and uses camouflage to exploit the color confusions of color deficient observers (Sloan & Habel, 1956; Belcher et al., 1958; Frey, 1958; Birch, 1997). The Ishihara test consists of single or double-digit numbers that have to be identified verbally and pathways for tracing for those who cannot read numbers. The 24-plate test version consists of the following: plate 1 for demonstration of visual task, plates 2-15 for screening, plates 16-17 for protan/deutan classification. The Ishihara test employs a range of designs such as transformation, vanishing, or hidden digit. In the vanishing type plate (Fig. 3, middle) a figure is seen by color normals but not by color deficients; the reverse of this, the hidden figure design, is harder to design and not always so effective. More complex patterns are contained in transformation plates (Fig. 3, left), with careful placement of the color dots giving an apparent transformation of the perceived figure; normal trichromats see one figure, and color deficient people see a different figure in the same design. Positive evidence of color deficiency is given by transformation designs, whereas vanishing designs give negative evidence. In the classification plate design (Fig. 3, right), protans only see the number on the right side of each plate and deutans only see the number on the left. The test is limited to red/green deficiency and cannot be used to assess loss of yellow-blue sensitivity.

The test is viewed at about two-thirds of one meter (arm's length) distance using a Macbeth easel lamp for illumination. The first 25 plates of the 38-plate test version were used in this investigation. The book is placed in the tray beneath the lamp and the illumination, equivalent to CIE Standard Illuminant C (representing average daylight), is incident at an angle of 45° to the plate surface. The illuminant used is important because the selected reflectances of the patches on the plates have been chosen with reference to this illuminant. The examiner instructs the person being tested to report the number they can see as the pages are turned and warns the subject that on some occasions they may not see a number. The first introductory plate is used to demonstrate the visual task. This plate is designed so that anyone, including color deficient subjects, should see this number. With a viewing time of about 4 seconds allowed for each plate, undue hesitation on the part of the subject is the first indication of color deficiency.

1.3.2 Dvorine Plate Test

The Dvorine test is based on pseudoisochromatic principles and is similar to the Ishihara test. It has 15 numeral plates, consisting of 1 initial plate that demonstrates the visual task, 12 plates for screening, and 2 plates for protan/deutan classification (see Fig. 4). These plates are of the vanishing type. The font of the numerals is slightly different to the Ishihara plates and is believed to be easier to read.The Dvorine test is administered in a similar manner to the Ishihara test. The plates are positioned at arm's length, perpendicular to the line of sight, under daylight illumination or a Macbeth easel lamp. The subject is instructed to read the numerals (all plates have a numeral).

Pseudoisochromatic plate tests provide a simple, readily available, inexpensive, and easy to administer screener mostly for red/green deficiencies. However, plate tests tend to be relatively easy to learn, and this encourages cheating. The spectral quality of the light source illuminating the plates is also important. Plates may be degraded by fingerprints, dust, and excessive light exposure. In general, subjects with minimal color vision loss tend to

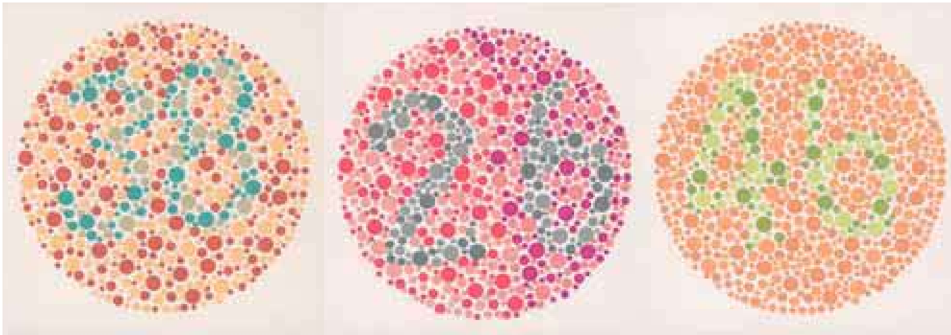

Figure 4: Dvorine pseudoisochromatic plates; left and right, vanishing design; middle, protan/deutan classification plate. Please note that the color of the plates may not be reproduced accurately in this document or in print since the printed reflectances and the viewing illuminant will be different.

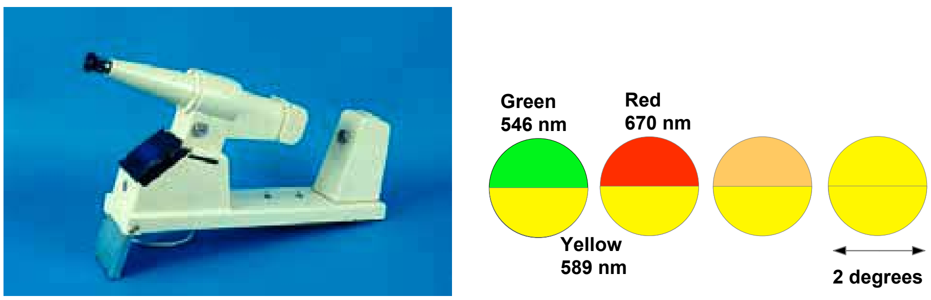

Figure 5: Photograph of the Nagel anomaloscope (Model I, Schmidt and Haensch, Germany) and illustration of the Nagel anomaloscope split field. The percentage mixture of red to green in the top half and the luminance of the yellow bottom field can be changed until a match of the two fields can be perceived.

show greater variability on repeated tests by comparison with normal trichromats and on some occasions can even pass these tests (Squire et al., 2005).

1.3.3 Nagel Anomaloscope

The Nagel anomaloscope (Fig. 5) is based on color matching and is the standard clinical reference test for identifying and diagnosing red/green color deficiency recommended by the National Research Council - National Academy of Sciences (NRC-NAS) Committee on Vision (1981). This instrument produces a disc stimulus that consists of two half fields and is viewed in an optical system. The top half of this disc is illuminated by a mixture of spectrally narrow red and green wavelengths, and the lower half is illuminated by spectrally narrow yellow light. Two control knobs are used, one to alter the red-green color mixture ratio of the top field, and the other to alter the luminance of the yellow lower field (see Fig. 5). The test is administered in two stages. Usually only one eye (i.e., the dominant eye) is fully tested and the other eye is then checked to ensure the same deficiency. This confirms that the loss of color vision is congenital. Following familiarization with the instrument controls, the subject is then asked to alter both the control knobs until the two halves of the circle match completely in both color and brightness. The subject is not asked to name the colors. A few matches are made, with the examiner "spoiling" the match after each setting. About ten seconds are allowed for each match and then, to minimize the effect of chromatic after images, the subject looks away from the instrument into the dimly lit room for a few seconds before the procedure is repeated. The second stage of the test is to determine the limits of the matching range. The initial matches made by the subject are used as a guide by the examiner to set the red/green mixture ratio near to the estimated limits of the range. The subject must alter the luminance of the lower yellow half of the field and see if an exact "match" in both color and brightness can be made with the set red/green mixture in the upper half. The ratio of the red/green mixture field is altered systematically by the examiner until the limits of the matching range are found. The matching range is recorded from the matching limits on the red/green mixture scale and the midpoint calculated.

Ideally, the red/green "match" parameters should provide enough information to determine whether a person has normal or defective red/green color vision, whether color deficiency is deutan or protan, and whether the subject is a dichromat (absence of a cone-type) or anomalous trichromat (anomalous cone-type). The size of the red-green matching range is often taken as an indicator of chromatic sensitivity loss. The red-green discrimination index (RGI), a parameter relating to the matching range, has been introduced in this study and provides an indication of the subject's ability to discriminate red-green color differences:

$$RGI = 1 - \frac{r_{subject}}{74}$$, where $r_{subject}$ is the subject's mean matching range.

The RGI ranges from a value very close to 1 for normal sensitivity, to 0 for a dichromat that accepts any red/green mixture setting as a match to the yellow field. A more appropriate measure of red/green sensitivity based on the Nagel is obtained simply by dividing the mean normal matching range (r_{mean}) obtained by averaging results for a large number of normal trichromats by the subject's range ($r_{subject}$). The mean normal matching range for the Nagel anomaloscope used in this study is approximately 4 scale units. Hence, the new measure of chromatic sensitivity becomes:

$$\text{Nagel sensitivity} = \frac{r_{mean}}{r_{subject}}$$

A scatter plot of Nagel midpoints on the red-green scale versus RGI allows one to separate a clear cluster of subjects with midpoints between 36 and 44 units on the red/green mixture scale, which are likely to be normal trichromats (see Fig. 6). Dichromats will accept the full range of red/green mixtures as a match with the yellow field (i.e. RGI=0), as they have only one photopigment

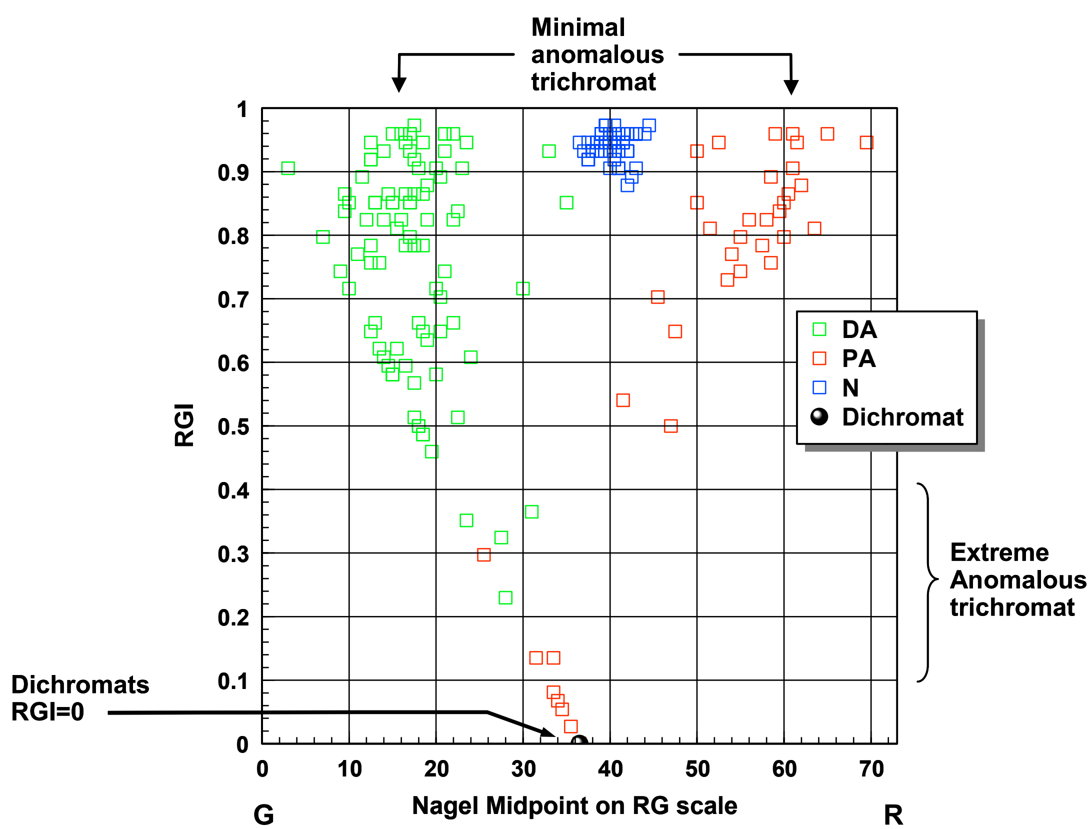

Figure 6: Scatter plot of matching midpoints versus RGI for 231 observers; 70 subjects formed a cluster that is separated from all other subjects by having a midpoint around 40 and a high RGI. The abbreviations in the legend refer to deuteranomalous (DA), protanomalous (PA), and normal trichromat (N) subjects. The value of the luminance setting on the yellow scale provides additional information to separate deutan and protan color deficient observers. The data show clearly that according to the Nagel test, many deutan and protan subjects have RG chromatic sensitivity that is indistinguishable from the range of values measured in normal trichromats.

in the spectral range provided by the instrument. Deuteranopes are distinguished from protanopes as the intensity of the yellow they set for both ends of the red/green scale is fairly similar, whereas protanopes set the luminance of the yellow very low to make a match at the red end of the scale and much higher at the green end. This is because protans tend to see red as fairly dark, as they have reduced long wavelength sensitivity. If a color match within the normal range is not achieved, the subject is classed as an anomalous trichromat. Two separate distributions are formed either side of the normal range, as protanomalous trichromats require significantly more red light in their color mixture, and deuteranomalous trichromats require more green (see Fig. 6). The RGI or matching range provides some measure of the severity of the color discrimination deficit on the Nagel anomaloscope, although it is well known that the correlation between the size of the matching range and the subject's ability to discriminate colors under more natural conditions is generally poor (Wright, 1946).

The principal advantage of the anomaloscope is that, unlike the previous tests, the parameters of the yellow match can be used to classify accurately the type of color deficiency involved.

1.3.4 Aviation Lights Test (ALT)

The Aviation Lights Test is a modified Farnsworth Lantern (Milburn & Mertens, 2004) designed to meet the FAA's signal color (USDOT-FAA, 1988) and International Civil Aviation Organization (ICAO, 1988) specifications for the red, green, and white signal light colors on aircraft. The chromaticity coordinates of the ALT are shown plotted in Fig. 11. Originally the test was employed for secondary screening of air traffic control specialist applicants.

Nine vertically separated pairs of colored lights (see Fig. 7) are presented to the examinee, who is seated 8 ft (~2.4 m) away from the lantern. The constant vertical separation of the 2 apertures is 13 mm, or 18.3 minutes of arc at the recommended viewing distance. Each pair of lights subtends a visual angle of 3 min arc. There are three colors: red, green, and white. Each series of nine pairs was presented three times in random order, making a total of 27 presentations.

Before the ALT test is carried out, the subject is shown each of the three test light colors. Light pairs numbered 1 and 2 are shown (see Fig. 7a) while saying: "this is green over red" and "this is white over green," respectively. The examinee has to name correctly the colors of the lights shown (with a pass criterion of not more than one error) in all 27 presentations. If the color of either or both lights in a pair was identified incorrectly, this was counted as one error. The ALT is administered in a very dim room that approximates the light level of the air traffic control (ATC) tower cab at night.

1.4 The CAD Test

The Color Assessment Diagnosis (CAD) test has been described in an earlier CAA report (CAA, 2006a). The CAD test is implemented on a calibrated visual display and consists of colored stimuli of precise chromaticity and saturation that are presented moving along each of the diagonal directions of a square foreground region made up of dynamic luminance contrast (LC) noise. The subject's task is to report the direction of motion of the color-defined stimulus by pressing one of four appropriate buttons. Randomly interleaved staircase procedures are used to adjust the strength of the color signals involved according to the subject's responses to determine the thresholds for color detection in each direction of interest to establish reliable estimates of red-green and yellow-blue color thresholds. The CAD test has a number of advantages over conventional tests, both in terms of isolation of color signals, as well as sensitivity and accuracy:

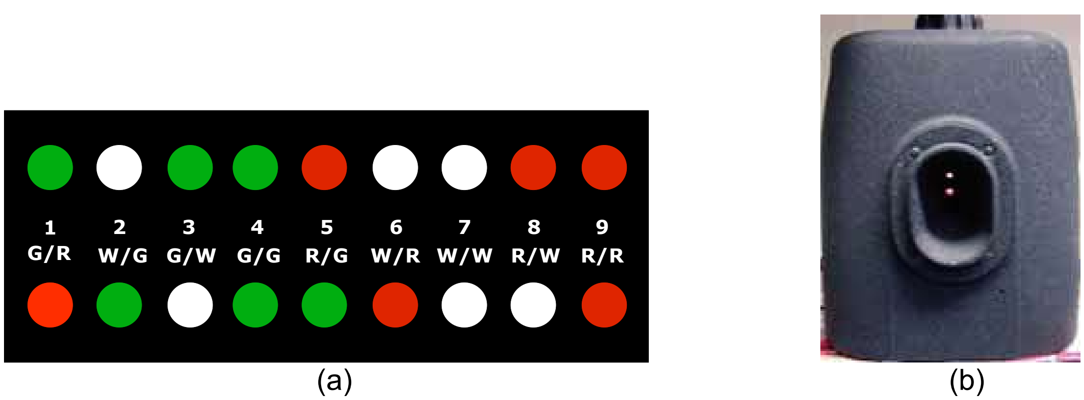

Figure 7: (a) Schematic representation of the different pairs of lights presented in the Aviation Lights Test (ALT). The lantern can show 9 different combinations (as shown) three times giving a total of 27 presentations, (b) Photograph of the ALT lantern.

- **Isolation of Color Signals**

It is very important to isolate the use of color signals by masking any luminance contrast cues. Although the colored stimuli generated are isoluminant for the standard CIE observer, the large variation in L:M spatial density ratio within normal trichromats (i.e., 0.6 to 13; Carroll et al., 2002) and the variation in cone spectral responsivity functions in color deficient observers introduce variations in the perceived luminance contrast of most colored stimuli. This is simply because the resulting luminance efficiency function V(λ) is likely to vary both among normal trichromats and within color deficient observers. The CAD test employs dynamic LC noise, and this effectively masks the detection of any residual luminance contrast signals that may be present in the colored test target. The mean luminance of the foreground remains unchanged, both spatially as well as in time, and equal to that of the surround background field. The technique isolates the use of color signals and ensures that the subject cannot make use of any residual LC signals. The dynamic LC noise does not affect the threshold for detection of color signals but effectively masks the detection of luminance contrast signals (Barbur et al., 1994; Barbur, 2004).

- **Measurement of Chromatic Detection Thresholds**

An efficient, four-alternative, forced-choice procedure is used to measure subject's chromatic detection thresholds in a number of carefully selected directions in the CIE – (x,y) chromaticity chart. Thresholds are measured along 16 interleaved directions in color space. These are grouped together to test red-green (RG) and yellow-blue (YB) color sensitivity. Threshold ellipses are computed and plotted using the standard CIE 1931 chromaticity chart. The use of 16 randomly interleaved color directions makes the technique statistically robust and eliminates any other possible cues, so the subject has to rely entirely on the use of color signals. The output of the CAD test also accurately diagnoses the class of color deficiency involved. If the latter is not needed, one only needs to test two color directions to screen for red/green color deficiency.

- **The Statistical Limits of Chromatic Sensitivity Within "Normal" Trichromats**

Chromatic discrimination thresholds have been measured in over 450 observers, including 250 normal trichromats and 200 color defective observers (Fig. 12). Fig. 8 shows the distribution of YB and RG chromatic thresholds obtained in the 250 normal trichromatic subjects. Fig. 9 shows the statistical limits for the "standard normal" (SN) observer on the CAD test plotted in the 1931 CIE-x,y color chart (Rodriguez-Carmona et al., 2005; Rodriguez-Carmona, 2006). The variability in both RG and YB thresholds is shown by the grey shaded ellipse, which represents the region of the CIE chart where we expect to find 95% of normal trichromats. The 2.5% and 97.5% limits define the boundaries of this region. The median chromatic discrimination threshold ellipse is also plotted. The median threshold value is important since it represents the Standard Normal (SN) observer. A subject's thresholds can then be expressed in SN units, and this makes it possible to assess the severity of color vision loss (i.e. an observer with a RG threshold of 2 SN units requires twice the color signal strength that is needed

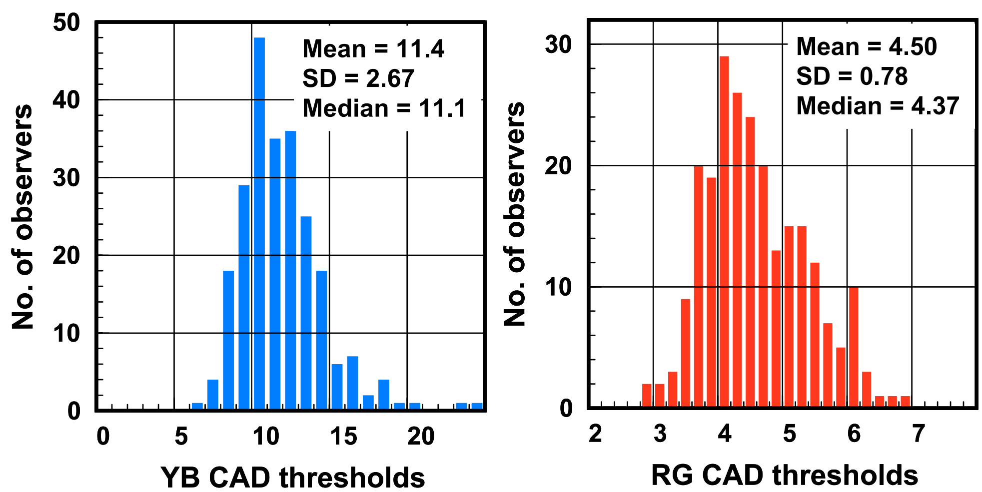

Figure 8: Frequency distributions of the YB and RG chromatic thresholds obtained in 250 observers with 'normal' trichromatic vision. The mean, standard deviation (SD) and median are shown.

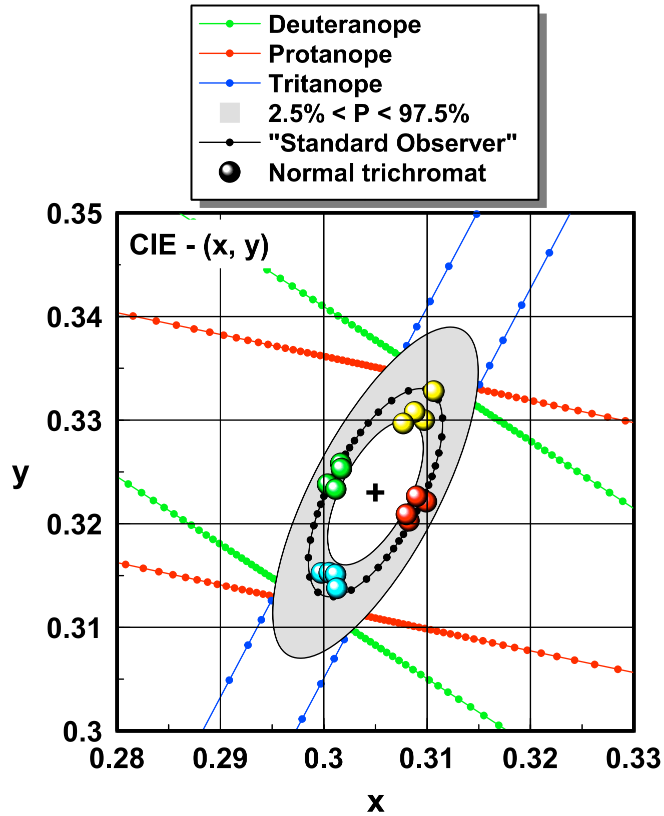

Figure 9: Data showing the 97.5 and 2.5% statistical limits that define the "standard" normal CAD test observer. The dotted, black ellipse is based on the median RG and YB thresholds measured in 250 observers. The grey shaded area shows the limits of variability of 95% of these observers. The deuteranopic, protanopic and tritanopic confusion bands are displayed in green, red and blue, respectively. The background chromaticity (x,y) is indicated by the black cross (0.305, 0.323). The colored symbols show data measured for a typical normal trichromat.

by the average standard CAD observer). Fig. 9 is an extremely useful representation in that it provides a CAD test template for the SN observer. Any subject's results provide instant diagnosis of either normal or deficient color vision when plotted on this template.

- **Detection of Color Vision Loss That Falls Outside Normal Range**

The distribution of thresholds along the directions examined provides enough information to classify even minimal deficiencies that would otherwise remain undetected using conventional color vision tests. For example, Fig. 10 shows results of two minimal color vision deficient observers that fall just outside the normal range indicated by the shaded grey area. The subject on the left (subject CC) has a Nagel range of 16-18 and passes the Ishihara, whereas the subject on the right (subject SH) has a Nagel range of 40-42 units but fails the Ishihara with 2 errors. Although both subjects are classified as minimum deuteranomalous on the CAD test, their classification on the other two tests is less clear. The first subject is classified as normal on Ishihara and deuteranomalous on the Nagel (but with a mixture range smaller than the average normal trichromat). The second subject is classified as normal on the Nagel, but slightly red/green deficient on the Ishihara.

- **Diagnosis of the Type of Deficiency Involved**

The CAD test identifies the type of deficiency involved by the elongation of the subject's results, either along the deuteranopic (Fig. 11, left) or protanopic (Fig. 11, right) confusion bands. In the case of absolute minimum deuteranomalous deficiencies, the distribution of the thresholds is as shown in Fig. 10. In the case of minimum protanomalous deficiencies, the thresholds are much larger and extend sufficiently along in the protanopic direction indicating a diagnosis of minimum protanomaly with no ambiguity. The agreement with the Nagel for screening and classification of the class of congenital red/green deficiency is ~99%.

- **Quantifying the Severity of Color Vision Loss**

The severity of red-green and yellow-blue loss of color vision is proportional to the color signal strength needed for threshold detection. For example, subjects in Fig. 11 show more severe loss (i.e., higher thresholds or lower chromatic sensitivity) than the subjects shown in Fig. 10. The severity of color vision loss can be quantified with respect to the standard normal observer (Fig. 8 and 9). Chromatic sensitivity varies greatly within color deficient observers from complete absence of red-green discrimination, in the case of dichromats, to almost normal sensitivity in subjects with thresholds not much larger than 2 SN units. Fig. 12 shows the subject's RG threshold in SN CAD units along the abscissa, plotted against the YB threshold along the ordinate in 450 observers. The results show that the RG thresholds vary almost continuously from very close to "normal" to extreme values which can be 25 times larger than the standard normal threshold. The YB thresholds, on the other hand, vary little, as expected, in the absence of yellow-blue loss or acquired deficiency. Interestingly, the RG thresholds show some correlation with YB thresholds in normal trichromats, suggesting that subjects with high RG chromatic sensitivity are also likely to exhibit high YB sensitivity. The loss of sensitivity (when expressed in Standard Normal (CAD) units (SN units) is greater in protanomalous than deuteranomalous observers (Fig. 12).

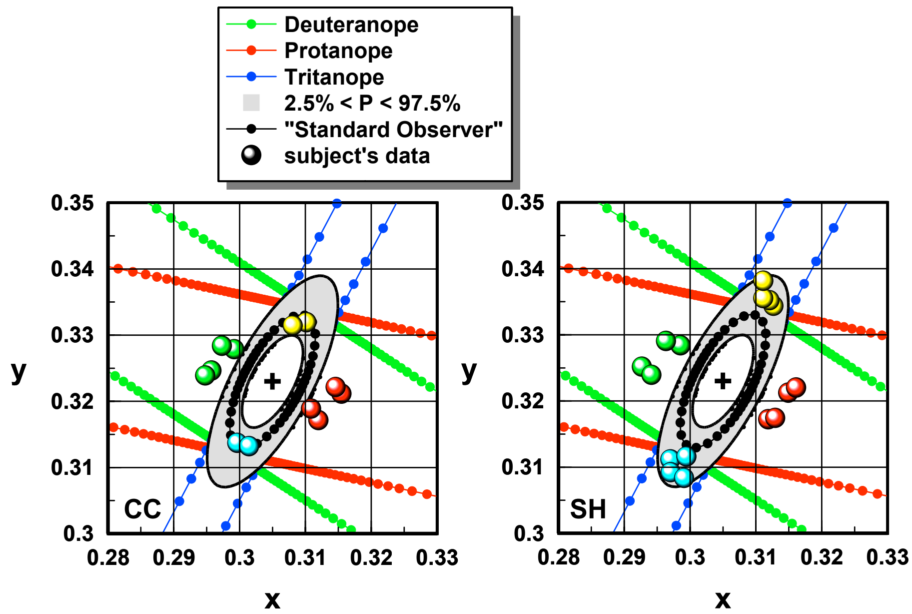

Figure 10: Chromatic thresholds for two color vision deficient observers with minimal color vision deficiency. The data for the average normal trichromat is shown as a black contour.

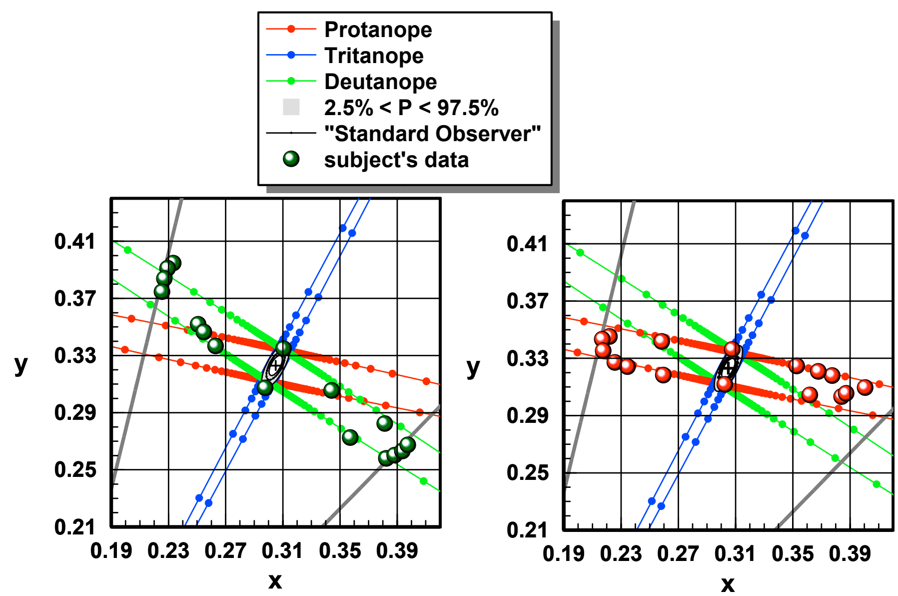

Figure 11: Chromatic thresholds for two color vision deficient observers with severe color vision deficiency. The largest chromatic displacements away from background chromaticity, as set by the isoluminant condition and the limits imposed by the phosphors of the display, are shown as grey lines. The extent of color vision loss is related to the elongation along the protanopic or the deuteranopic confusion band and suggests that the greater the elongation, the lower the level of chromatic sensitivity.

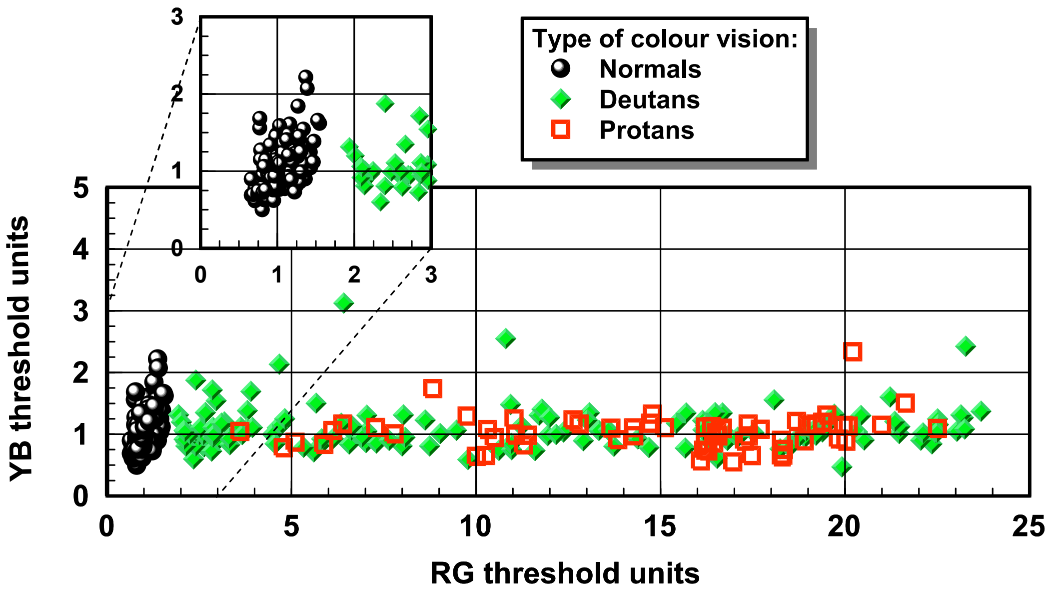

Figure 12: Graph showing Red-Green (RG) and Yellow-Blue (YB) thresholds expressed in CAD Standard Normal units for the population of subjects tested as part of this study. The spread of the data along the abscissa illustrates the large variation that exists among subjects with deutan- and protan-like deficiencies.

- **Effects of Light Level and Stimulus Size**

Both the ambient light adaptation level and the size of the colored stimulus can affect chromatic sensitivity. In general, as the light level is reduced and/or the stimulus size is decreased, the RG and YB thresholds increase. The YB thresholds are affected most at lower light levels. Both background luminance and stimulus size have been optimized for the CAD test so that no significant improvement in chromatic sensitivity results by increasing either the light level or stimulus size. Any small variations in either light level or stimulus size will not therefore alter significantly the subject's RG and YB thresholds (Barbur et al., 2006). However, older subjects are likely to show more rapid effects as the light level is reduced because the retinal illuminance in these subjects is already low as a result of small pupil sizes and increased pre-receptoral absorption of blue light.

- **The Effects of Aging on Red-Green and Yellow-Blue Loss of Chromatic Sensitivity**

The effect of aging for YB and RG chromatic thresholds in normal trichromats is shown in Fig. 13. These results show that up to the age of 60 years there is little correlation between the subject's age and chromatic sensitivity. A small effect can be observed when examining YB thresholds (but the correlation with age remains very poor) and virtually absent in the case of RG thresholds. The age range examined is representative of the typical working life of pilots. Color vision is usually assessed in demanding occupational environments. Loss of color vision later in life is described as acquired color deficiency and can be caused by a number of factors including both systemic diseases and specific diseases of the eye (such as diabetes, glaucoma, age-related macular degeneration.). Since loss of chromatic sensitivity usually precedes the reliable detection of any structural changes using fundus imaging, regular screening for acquired color vision loss may be of great clinical value. In view of these findings, it makes sense to recommend that in addition to assessing color vision at the start of the working career, periodic re-assessments should also be done as a way of testing for acquired deficiencies.

1.5 Summary of Congenital Color Vision Deficiencies

Congenital color vision deficiencies remain unchanged throughout life and are largely determined by changes in the spectral responsivity of cone photoreceptors that are determined genetically. There are a number of other factors that can affect chromatic sensitivity, such as the optical density of photoreceptors, post-receptoral amplification of cone signals, or pre-receptoral filters that are spectrally selective and reduce the amount of light that reaches the cone photoreceptors in the eye (Alpern & Pugh, Jr., 1977; Alpern, 1979; Neitz & Jacobs, 1986; Barbur, 2003). These factors are all likely to contribute to the variability measured within normal and color deficient observers.

Congenital yellow/blue color vision deficiency is very rare (with an incidence of 1 in 13000 to 65000; Sharpe

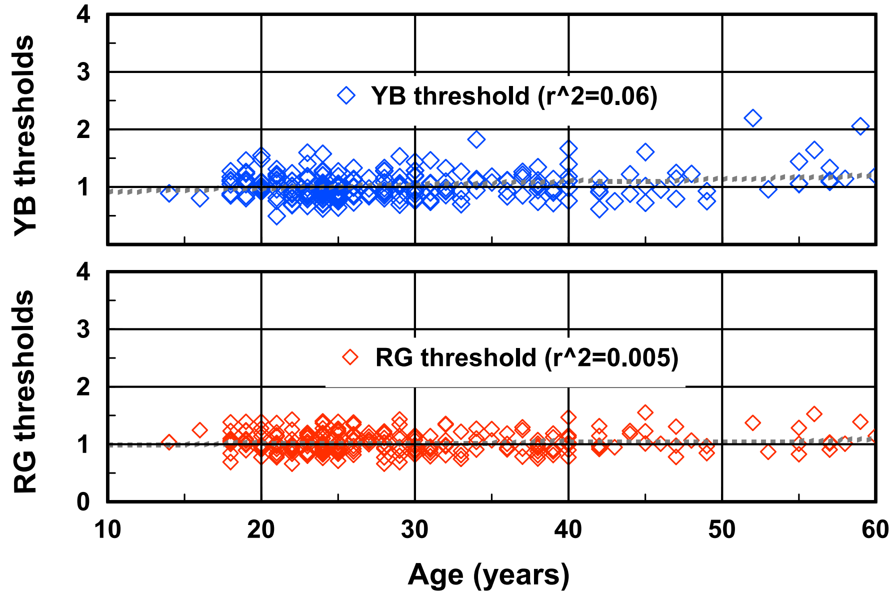

Figure 13: Effect of age on the YB and RG chromatic thresholds for normal trichromats under normal daylight conditions. The best-fit line is shown for both sets of data. The correlation coefficients (r^2) for the YB and RG thresholds are 0.06 and 0.005, respectively.

et al., 1999) and usually implies the absence of S-cones. Loss of YB sensitivity with age, on the other hand, is very common and is often associated with toxicity or disease (see below).

1.6 Acquired Color Vision Deficiencies

Acquired deficiencies tend to affect both RG and YB color discrimination, although frequently the YB loss is greater and more apparent. Acquired color deficiencies are most commonly caused by systemic diseases (e.g., diabetes, multiple sclerosis) and other diseases that are specific to the eye (e.g., glaucoma, age-related macular degeneration, optic neuritis). Acquired deficiency affects predominantly older subjects. Acquired loss can sometimes be expressed in subjects with congenital color deficiencies. If congenital color deficiency is present, the identification of acquired color deficiency and the classification of the congenital component are more difficult. In such cases, the use of larger stimuli and dynamic luminance contrast noise that achieves a high level of luminance contrast masking with saturated colors can reveal both the type of congenital deficiency and the acquired loss of chromatic sensitivity (Barbur et al., 1997). There are other differences as well. Acquired loss of color sensitivity is generally non-uniform over the retina in the same eye and often affects one eye more severely than the other. One can also separate the congenital and the acquired loss by carrying out the CAD test in each eye both at the fovea and in the near periphery of the visual field or/and by using more than one stimulus size. The congenital component remains largely unchanged, whereas the acquired component varies with stimulus size, retinal location, and eye tested. Since the yellow-blue sensitivity is most affected, the CAD is particularly suitable for investigating acquired deficiency since it tests for both red-green and yellow-blue loss. Fig. 14 shows examples of acquired color vision deficiencies.

2.0 SUBJECTS AND METHODS

Summary of tests employed in this study:

1. Ishihara plate test
2. Dvorine plate test
3. Nagel anomaloscope
4. CAD test
5. Aviation Lights Test (ALT)
6. PAPI simulator test
7. PAPI Signal Lights test (PSL)

The PAPI and PSL simulators were designed and constructed specifically for this investigation. A full assessment of color vision using all these tests takes between 1.5 to 2 hours per subject. The order in which the different tests were carried out varied randomly, and the testing took place in three different rooms, allowing the subject to take short breaks between tests. We examined 182 subjects in this investigation: 65 normal trichromats and 117 subjects with deutan- and protan-like color deficiencies. The age of the subjects ranged from 15 to 55 years (mean 30.2 years, median 27 years).

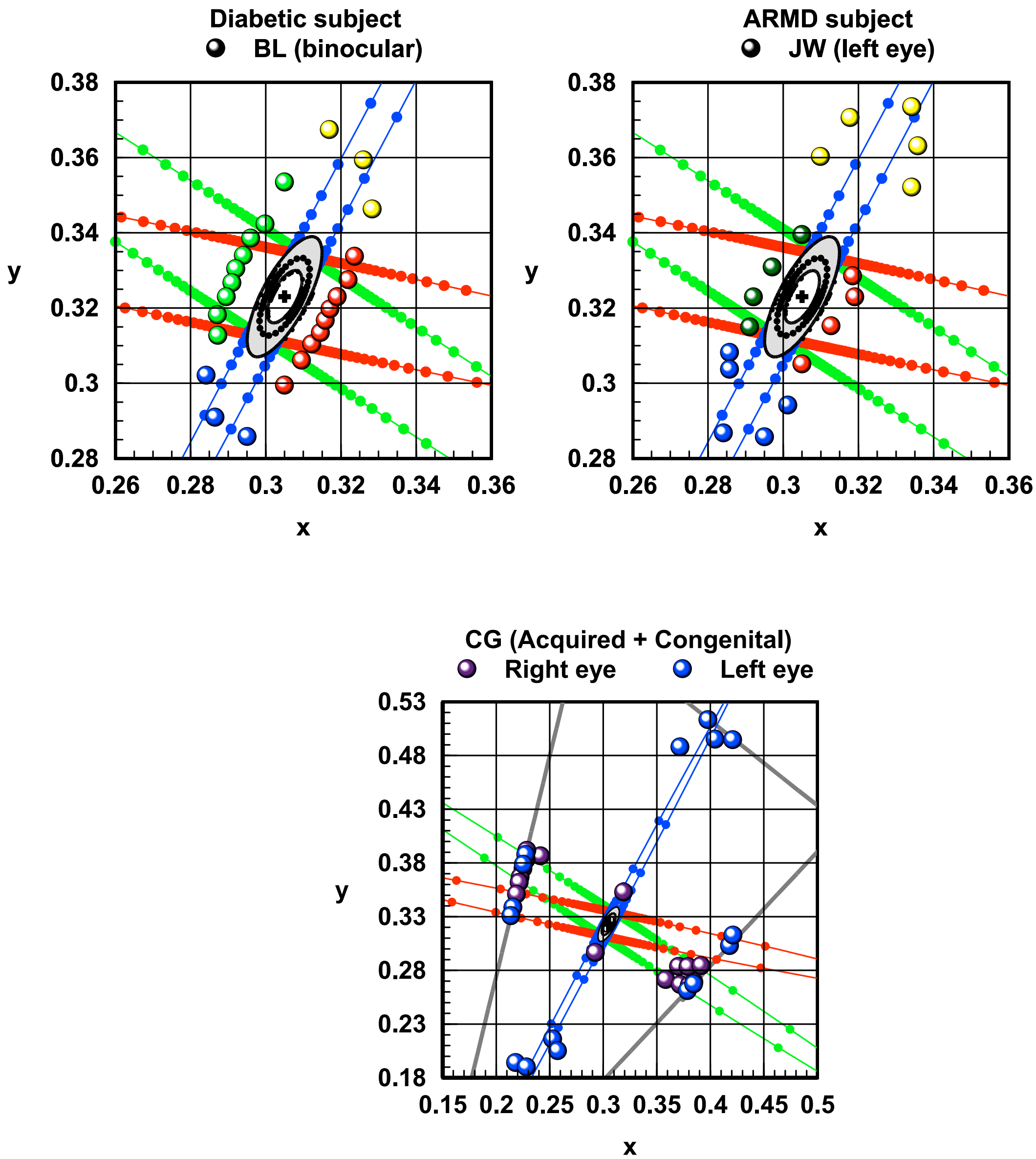

Figure 14: Examples of subjects with acquired loss of chromatic sensitivity. The data shown on the top left graph is from a subject with diabetes; top right shows data from the left eye from a subject with Age Related Macular Degeneration (ARMD); and bottom graph shows data from a subject with both congenital and acquired color vision loss (note the difference between the two eyes).

2.1 PAPI and PSL Simulator

The PAPI system consists of four lights arranged in a horizontal line and installed at right angles to the runway with the nearest light some 15 m away from the edge. The lights are approximately 30 cm in diameter with an inter-light separation of 9 m. The unit nearest the runway is set higher than the required approach angle at 3°30', with progressive reductions of ~20 minutes of arc further out field: 3°10', 2°50' and 2°30' (for a 3° approach). Usually each unit contains three light projectors (in case one fails). The light system has an intensity control for day and night use, with up to six luminous intensity settings: 100%, 80%, 30%, 10%, 3%, and 1% (CAA, 2004).

The units direct a beam of light, red in the lower half and white in the upper half, towards the approach. The different elevation angles give a combination of red and white for an on-slope signal, all-red if the aircraft is too low, and all-white if it is too high (see Fig. 1). The chromaticities of the lights should follow the ICAO specification for Aerodrome Ground Lighting (AGL) (see Fig. 15). The light intensity of the white signal is required to be no less than twice and no more than 6.5 times as bright as the red signal. The recommended intensities for the white and red light are 85000cd and 12750cd, respectively, at the maximum of their light intensity distribution (CAA, 2004).

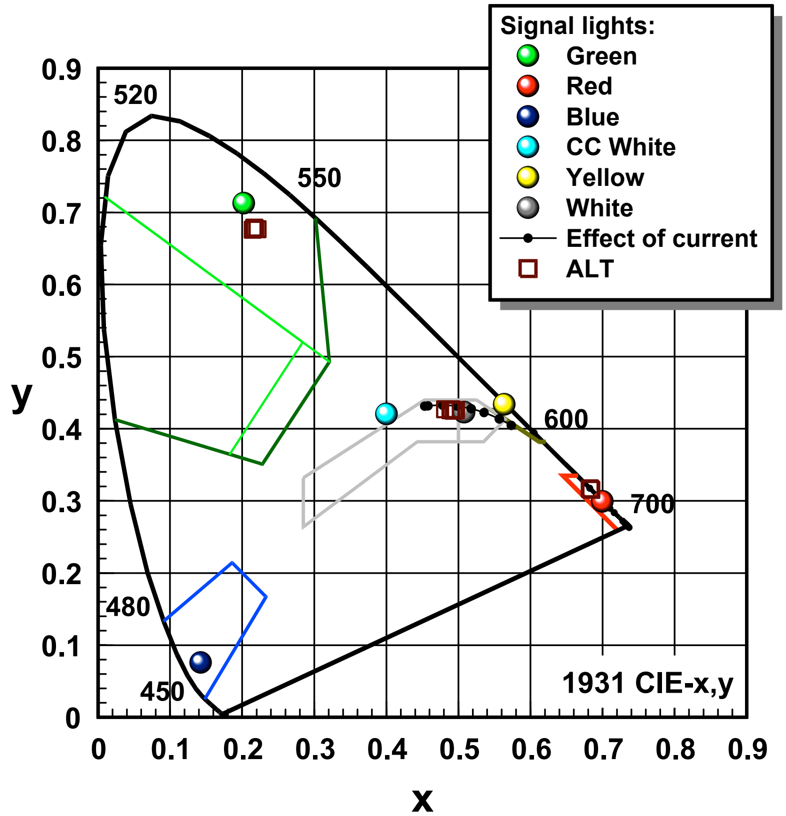

Figure 15: 1931 CIE-x,y color space diagram showing the recommended chromaticity boundaries for the colors of light signals (CIE, 2001a; CIE, 2001b). The signal colors used in the laboratory set-up are also plotted and the effect of varying the current (i.e., the output intensity of the laps) is shown for the white light.

A schematic of the laboratory set-up developed to simulate the PAPI and PSL tests is shown in Fig. 16. A four-channel optical system was developed using an airfield halogen lamp (JF6.6A100W/PK30d) as the single light source. The light is then split up and channeled using two beam-splitters (BS) to generate four beams. Each beam of light passes through two motorized filter wheels: color (CW) and neutral density (NDW) wheels. The CW has six different filters: red, modified white (~3900K), blue, green, yellow, and standard white (~2400K). Each NDW has neutral density filters with the following optical density (OD) values: 0.0, 0.3, 0.6, 1.0, 1.3, and 1.6. During the calibration procedure, the luminous intensity of each beam was measured with each filter in place to account for the actual and not the nominal absorption of each filter. The angular subtense of each light was 1.36' at a viewing distance of 4 m. Beyond ~0.8 km the angular subtense of the real PAPI lights approaches the diffraction limit of the eye. The size of each light on the retina remains relatively unchanged as the approach distance is increased, but the light flux captured from each light is decreased. On approach, the PAPI lights are first seen as a small continuous line until the angular separation between adjacent lights is resolved by the eye (typically less than 2', taking into consideration pupil size and optical aberrations). In order to reproduce the geometry of the real PAPI lights in the laboratory for a viewing distance of 4 m, the adjacent lights (center to center) were separated by ~6.5 mm. This corresponds to an angular separation of 5.5' which translates to an approach distance of 5.54 km, in the case of the real PAPI lights. This design therefore requires the pilot to locate and recognize the white and red PAPI lights from 5.54 km when the size of the image of each light on the retina is determined by the point-spread-function (PSF) of the eye. We did not choose a longer approach distance in order to minimize the effects that higher order aberrations and increased scatter in the eye have on the retinal images of the lights. When the pupil size is large, the higher order aberrations in the eye can be quite large, and this causes the PSF to broaden and the visual acuity to decrease. The light distribution in adjacent PAPI lights can overlap significantly, and this, in turn, makes it more difficult for the subject to process the color of each light. Since a longer approach

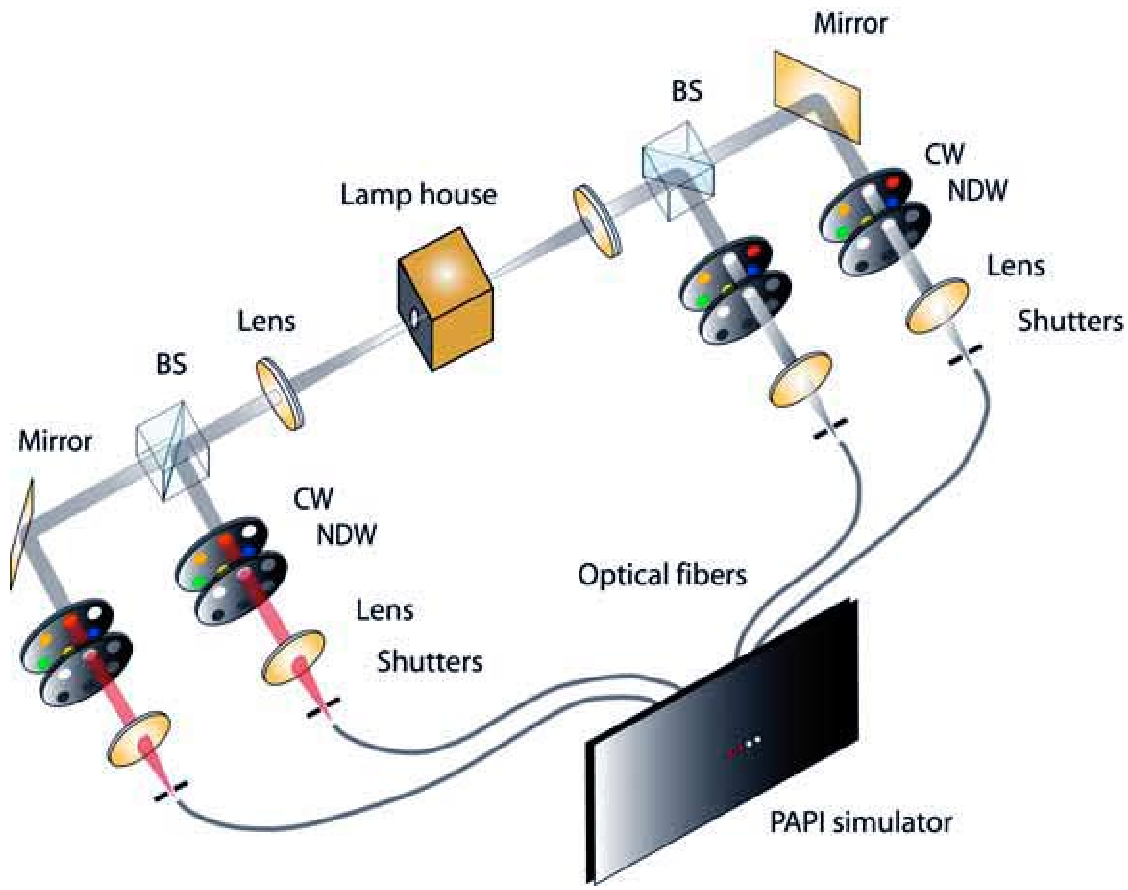

Figure 16: Schematic representation of PAPI simulator designed and constructed for this study. Light emerging from each arm of the lamp house is divided into two channels via beam splitters (BS) to produce four independent channels. Each channel has a color wheel (CW) and a neutral density wheel (NDW) which are controlled by the computer. After passing through appropriate filters, the light from each channel is focused into an optical fiber head which are attached to the viewing panel so as to simulate the PAPI lights.

distance would produce even more overlap, a distance of 5.54 km, which is considered to be safe, was selected for the study.

The optical fiber heads form a line located at the center of a black plate, which provides a dark uniform surround (see Fig. 16). The whole system is encased and ventilated by two fans to prevent overheating and to reach a steady state temperature which is needed for stable lamp operation. The intensities of the red and white lights were adjusted using ND filters so that the simulated PAPI lights appeared as intense as the real PAPI when viewed in the dark from a distance of 5.54 km. These calculations assumed that in the absence of significant ambient light, the pupil of the eye would in general be large (~6 mm). In addition, the intensities of the colored lights varied randomly by ±0.3 OD with respect to the nominal values to eliminate the detection of brightness cues.

The effect of the different intensity settings and ND filters was investigated to establish the extent to which the chromaticities of the white and red lights changed with lamp current setting and/or the use of ND filters (see Fig. 17). The results show that the ND filters caused only small changes in the chromaticity coordinates of the white, and even less for the red light. Changes in lamp current caused larger changes in the chromaticity of the white light, but in spite of these changes the white remained within the "variable white" area indicated on the CIE diagram, as appropriate for AGL (see Fig. 15 and 17). In the case of real PAPI lights, other factors, such as atmospheric absorption, can also affect the chromaticity of the white, with very little effect on the red.

2.2 Testing Procedure

The four horizontal lights are presented for 3 seconds and the subject's task is to simply report the number of red lights in the display. There are five possible combinations of red and white lights that are presented randomly (Fig. 18, left). When carrying out the PAPI simulator test,

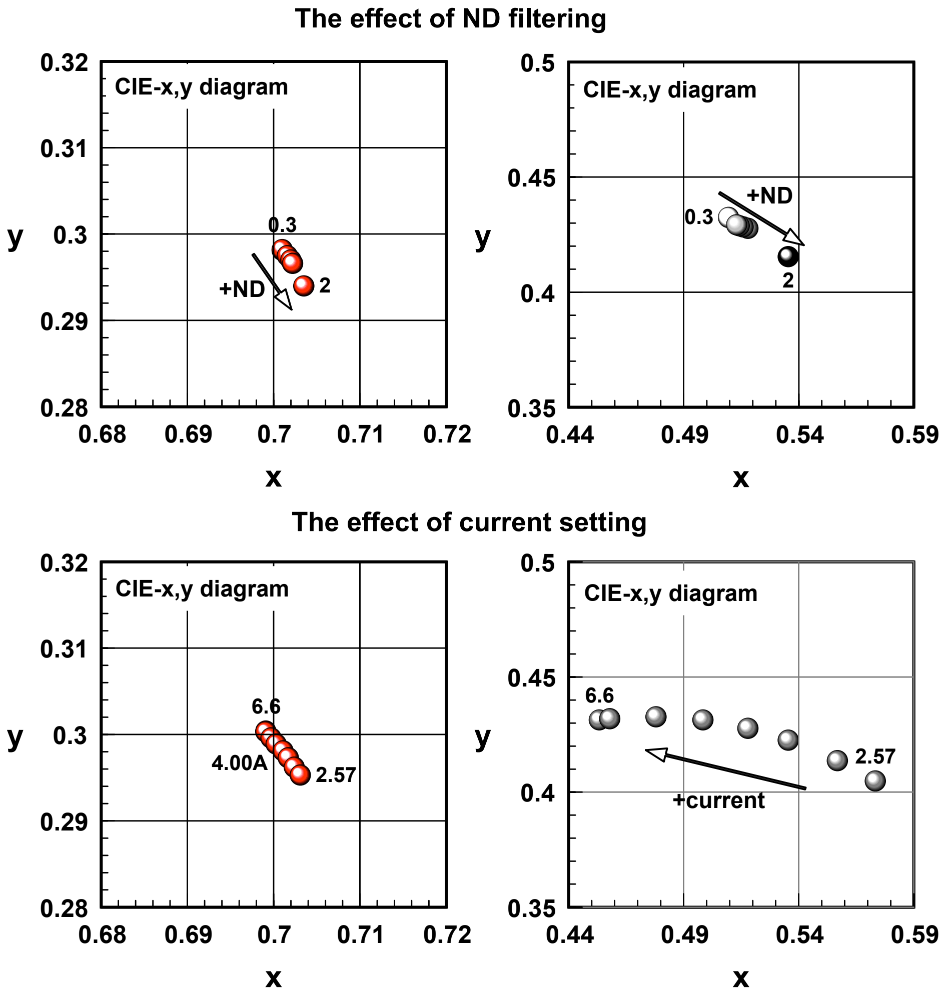

Figure 17: Graphs showing the CIE-x,y chromaticity coordinates of the Red and White lights under the effect of neutral density (ND) filtering and the effect of the current setting (intensity) on the lamp in amps (A).

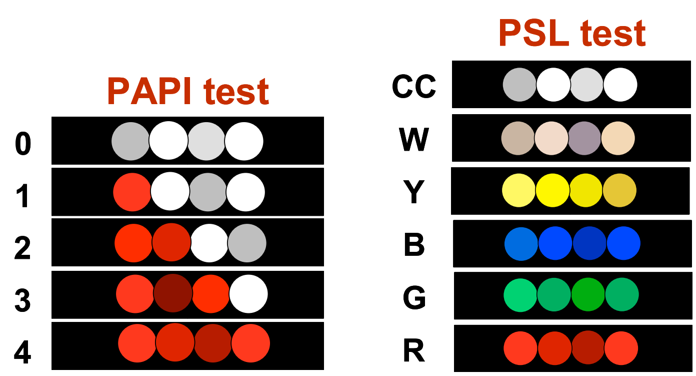

Figure 18: Schematic representation of the Precision Approach Path Indicator (PAPI) simulator test (left) and PAPI Signal Lights (PSL) test (right). The PAPI test presents 5 different conditions (as shown) twelve times giving a total of 60 presentations. The PSL presents 6 different conditions (as shown) twelve times giving a total of 72 presentations. The intensities of the colored lights varied randomly by ±0.3 OD with respect to the nominal values to eliminate the detection of brightness cues.

observers were instructed to report the number of red lights using the following names: one, two, three, four, or zero (to avoid confusing "none" with "one"). Prior to the test, observers were allowed to dark adapt to the low mesopic surround and then were presented with a practice run. A low power lamp was placed behind the test equipment to provide low mesopic conditions of ambient illumination. The black, immediate surround around the PAPI lights, was dark (i.e., mean luminance ~0.005 cd/m2). Subjects were encouraged to respond only after an auditory cue signaled the end of the 3-second viewing period. The PAPI test was carried out twice, once with the standard white (~2400K) and once with a modified white light (higher color temperature of ~3900K).

The PSL uses similar parameters to the PAPI lights test. In this test six possible colors are presented (standard white, modified white, red, green, blue, and yellow). The chromatic properties of the lights lie within the boundaries for the recommended signal lights for AGL (CAA, 2004) as shown in Fig. 15. The PSL addresses the issue of correct color naming when all lights have the same chromaticity, as opposed to the ability to distinguish and categorize some of the four lights as red and the others as white on the basis of some perceived differences between the lights. The PSL tests whether the applicant can recognize and name reds as "red" and whites as "white' for the same conditions and geometry as the PAPI lights, but when all the lights are of the same color. The conditions when all four PAPI lights have the same color to indicate "far too low" (all reds) or "far too high" (all whites) are clearly very important. Observers were instructed to report the color of the lights as either: red, green, yellow, blue, or white.

There were two whites: the standard white, as produced by the lamp, and a modified white, produced by raising the color temperature of the standard white by 200 MIREDS. This is achieved by using a color correction filter that modifies the spectral content of the tungsten light to make it more like daylight. Prior to the test, observers were presented with a practice run. All the colors were shown to the subject and named by the examiner during the practice run, and the subject was allowed to review any of the lights and to ask the examiner to confirm the color. The results for the PAPI and PSL are recorded as the percent correct.

3.0 RESULTS

The color vision of each subject was examined using five different color vision tests, as well as the PAPI and PSL simulator tests. Results from each of the five tests were then examined, in relation to the subject's performance on the PAPI, to establish which test yields the best prediction of performance in the PAPI task. Performance on the PAPI task is computed as number of correct reports out of a total of 60 presentations.

The results summarized in Fig. 19 show that normal trichromats can also make errors, both on the PAPI and the Ishihara tests (i.e., five subjects produce one error, one subject produces two errors, and one other subject produces three errors on the PAPI). The rest of the normal

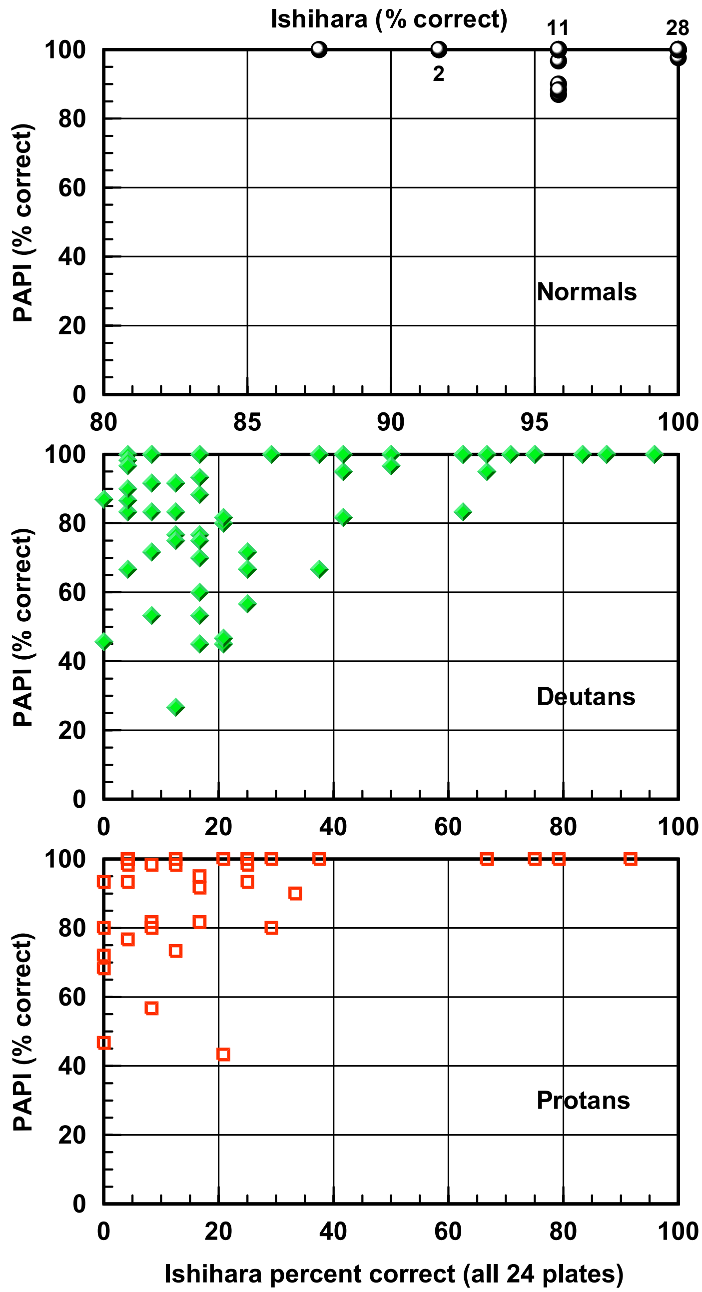

Figure 19: The number of plates read correctly expressed as a percentage on the Ishihara test (24 plates) is compared to performance on the PAPI simulator test separately for normals, deutan and protan color vision observers. The x-axis for the top graph has been expanded to show clearer the errors made by normals.

subjects score 100% correct on the PAPI test. Results for deutan color deficient observers reveal that all subjects with scores >70% (i.e., 16 or more correct plates out of 24 on the Ishihara 24-plate test) pass the PAPI with a score of 100% correct. Results for protan observers show that four subjects with scores greater than 40% pass the PAPI test. Overall, the results reveal very poor correlation between the subjects' performance on the Ishihara and the PAPI test scores. Many of the subjects that pass the PAPI task can score anything from 0 to 95% correct on the Ishihara test.

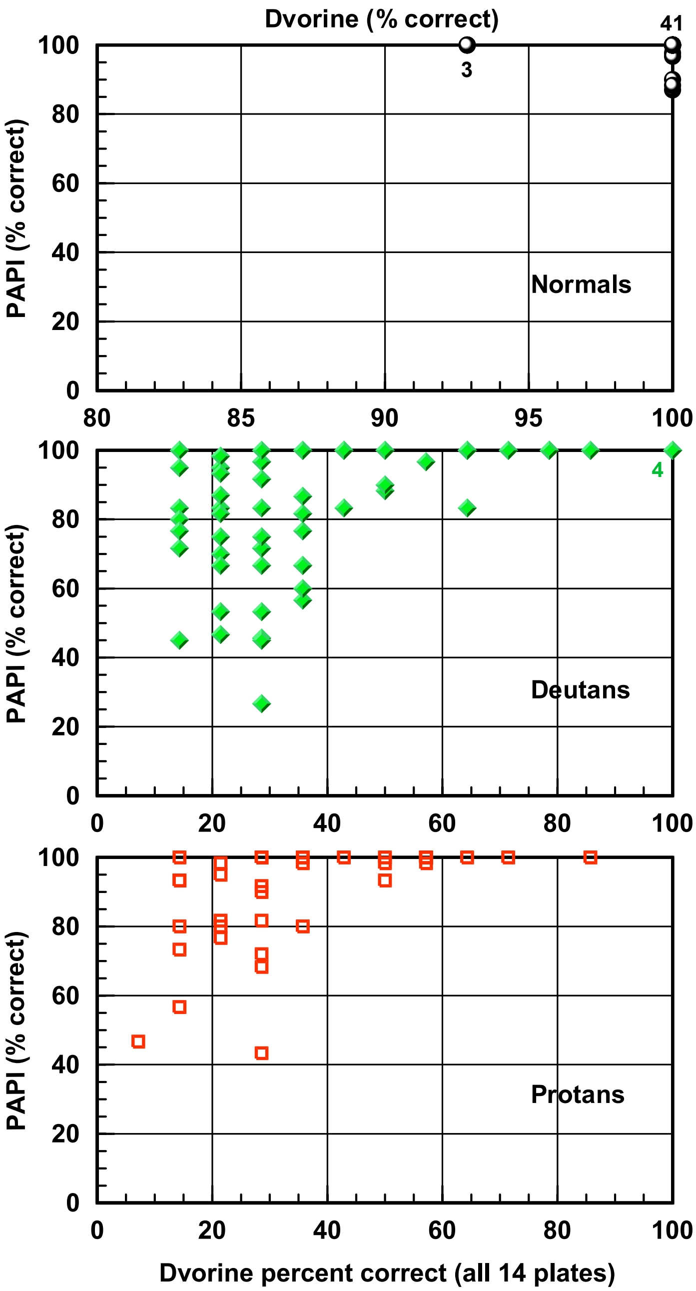

Figure 20: PAPI percent correct scores plotted as a function of the number of plates read correctly on the Dvorine test (expressed as a percentage) for normals, deutan and protan observers.

Comparisons of data from the Dvorine plate test with the PAPI simulator show similar results to those obtained with the Ishihara test (Fig. 20). Three normals obtain less than 100% on the Dvorine test (but pass the PAPI with no errors). Deutan and protan color deficient subjects need more than 65 and 50%, respectively, on the Dvorine plate test to achieve 100% on PAPI. Since the prediction of the class of deficiency (protan or deutan) involved is poor with both Ishihara and Dvorine tests, it is difficult to know which of the two limits one should apply to any color deficient subject.

Fig. 21 plots the PAPI scores against the subjects' performance on the ALT test. All normals secure 100% score on the ALT, but not on the PAPI test. Results for deutanobservers show that out of 77 subjects tested, only 18 passed the ALT (with a pass criterion of one or no errors). Fourteen of the 18 subjects that passed the ALT also passed the PAPI. Results for protans show that all 40 subjects tested had failed the ALT test and that only one subject achieved a score higher than 80%, although just over 50% of protans passed the PAPI.

Fig. 22 compares PAPI scores with a measure of RG sensitivity based on the Nagel anomaloscope range. Only a few deutan and protan observers pass the PAPI with Nagel sensitivity >0.6 (deutan) and >0.4 (protan). The Nagel anomaloscope test is excellent at distinguishing between deutan- and protan- like deficiencies, but fails to quantify reliably the severity of color vision loss, and does not test for yellow/blue deficiency.

PAPI test scores in Fig. 23 are plotted against the corresponding CAD based measure of RG sensitivity. The top section shows the performance in normal trichromats. Most normal trichromats perform 100% correct in the identification of the red and white lights. However, 7 out of the 65 normal subjects made some errors. This could be due to lack of attention and/or reduced visual acuity at low light levels caused by increased higher order aberrations when the pupil size is large. The errors were found to be distributed with equal probability among the five conditions. The blue dotted line in Fig. 23 shows the 95% confidence interval. The lower sections compare the performance on the PAPI test with the corresponding CAD measure of RG sensitivity for subjects with deutan- and protan-like deficiencies. The sensitivity limits beyond which deutans and protans perform the PAPI test, as well as normal trichromats, are 0.17 (RG threshold ~6 SN units) and 0.085 (RG threshold ~12 SN units), respectively. These limits are indicated by dotted vertical lines in Fig. 23.

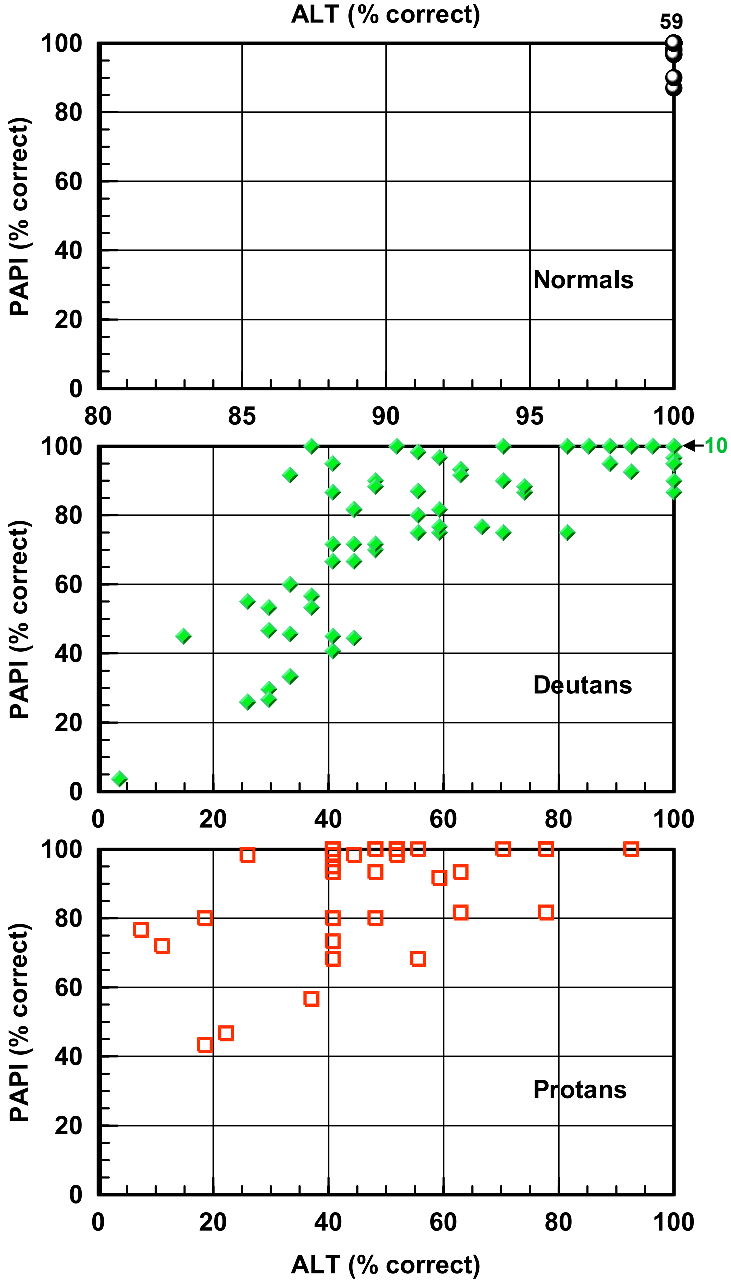

Figure 21: The number of presentations identified correctly on the Aviation Light Test (ALT) from a total of 27 presentations is compared to performance on the PAPI simulator test separately for normals, deutan and protan color vision observers.

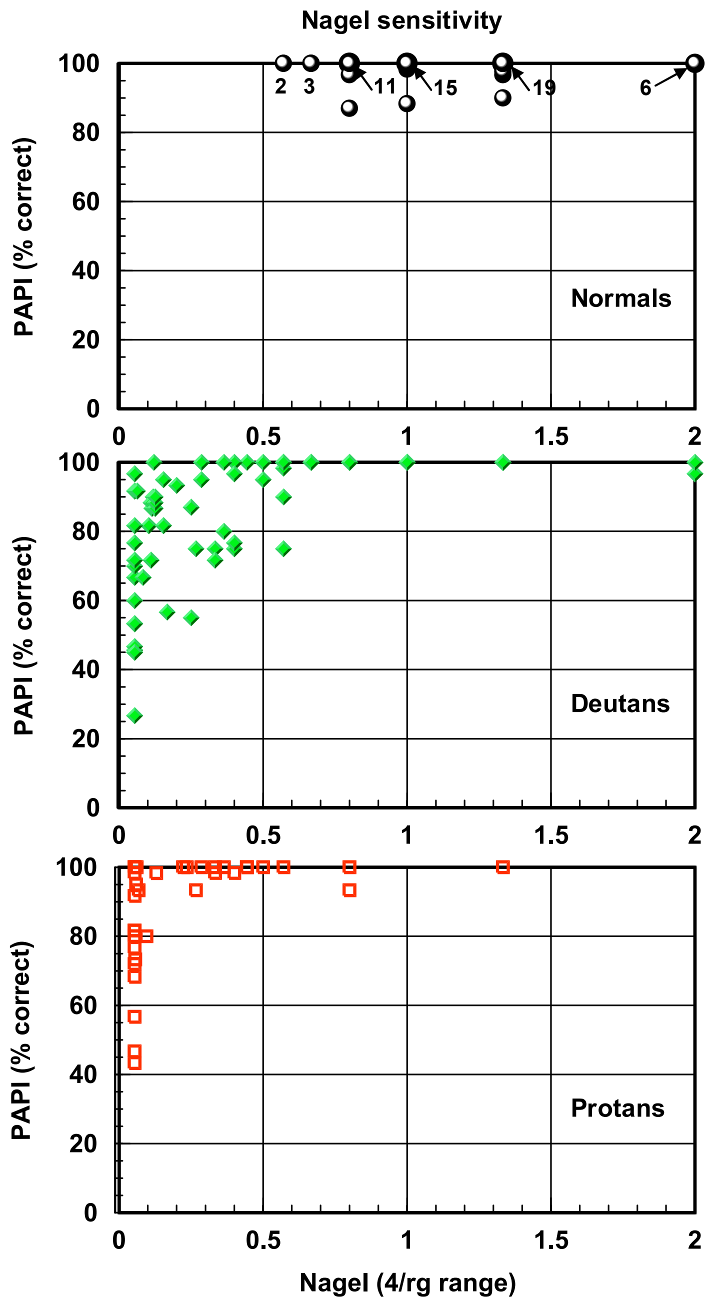

Figure 22: PAPI test scores plotted against an index of red-green chromatic sensitivity based on the Nagel anomaloscope range. Results are again shown separately for normals, deutan and protan observers. Numbers next to some symbols indicate number of subjects with overlapping data points.

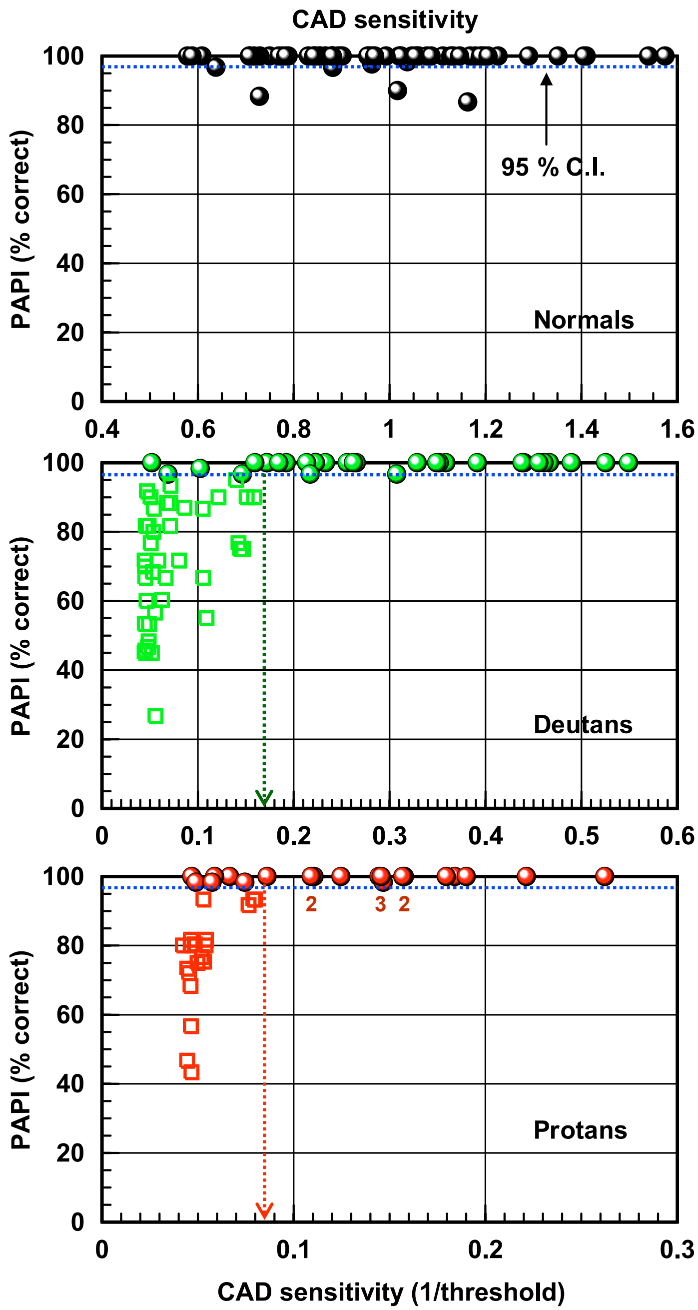

Figure 23: Graphs showing performance of normal, deutans and protan observers on the PAPI (standard white) versus CAD test sensitivity (1/threshold). Note the scale of the x-axis is different for the three graphs for clarity of presentation.

Fig. 24 shows the benefit of replacing the standard white light in the PAPI test with the color-corrected (CC) white. Higher PAPI performance scores were obtained with the CC white in all subject groups. All but one normal trichromat scored 100% correct on the PAPI test when using the CC white. Fig. 24 shows the overall improvement observed in the deutan group. All deutans obtain 80% correct or higher, and many score 100% correct with the CC white. The improvement is, however, very small among protanomalous observers, particularly among subjects with severe deficiency (i.e. those with RG sensitivity less than 0.05 units). The improvements were statistically significant in normal trichromats ($p=0.019$) and deutans ($p=0.000$). Protans also showed slightly higher PAPI performance scores with the CC white, but the improvement failed to reach statistical significance ($p=0.224$).

The PSL test was introduced to examine the condition when all four PAPI lights have the same color. In this condition the subjects can no longer make use of perceived differences between two adjacent colors presented simultaneously without being able to name the correct color. The results in Fig. 25 show that subjects do not often confuse reds with whites or whites with reds. Four protan subjects with severe loss of chromatic sensitivity (i.e., CAD RG sensitivity less than 0.05 units) make W=R and R=W errors, and six deutan subjects with CAD RG sensitivity less than 0.15 units make errors on W=R.

3.1 Computing an Index of Overall Chromatic Sensitivity

The threshold signal needed to just see the colored stimulus is expressed by direct comparison with the median threshold for normal trichromats. This approach has the advantage that a threshold <1 indicates color discrimination better than the standard normal trichromat, whereas values >1 indicate precisely the increase in threshold signal with respect to the normal observer. Sensitivity is usually

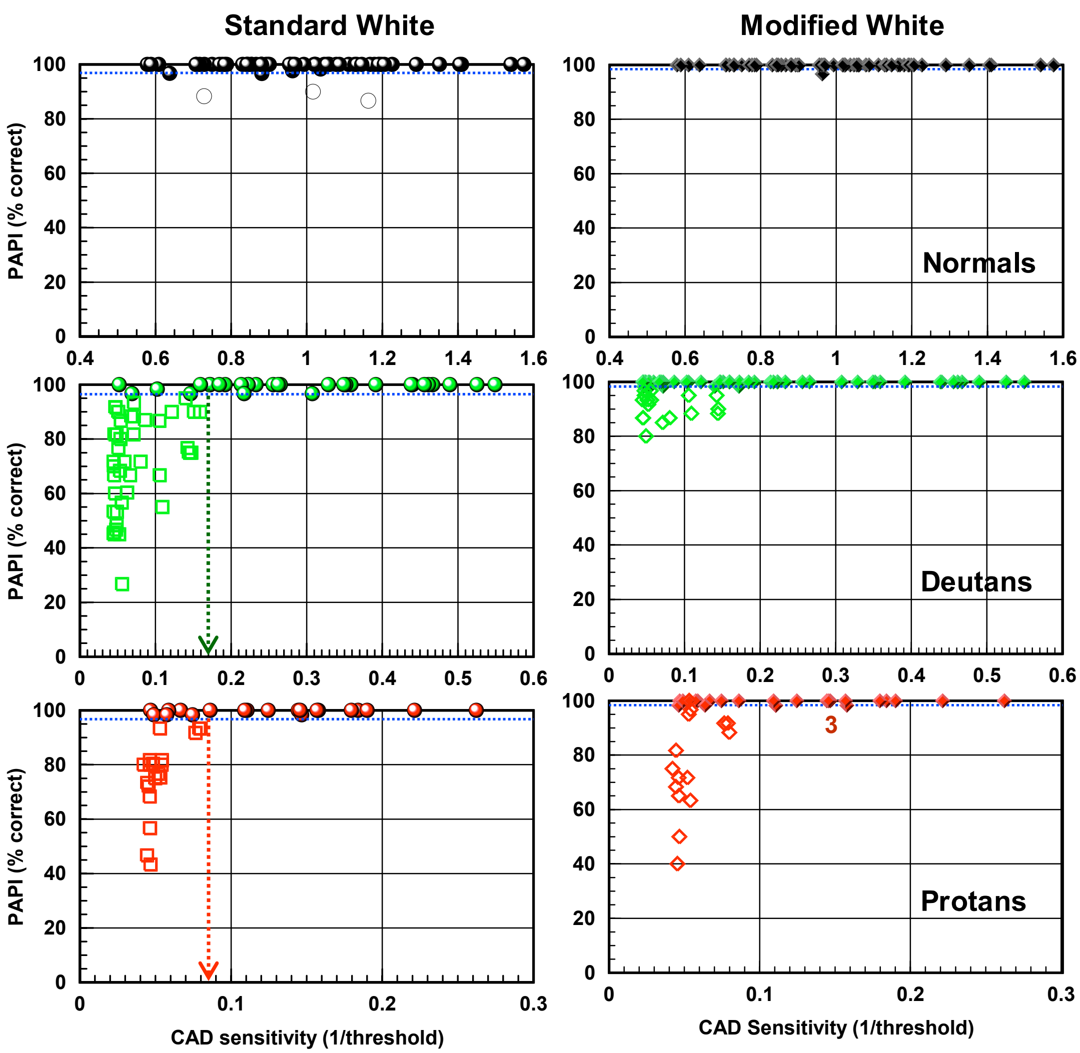

Figure 24: Graphs showing comparisons between standard and modified white (color-corrected white) versus the CAD test sensitivity values.

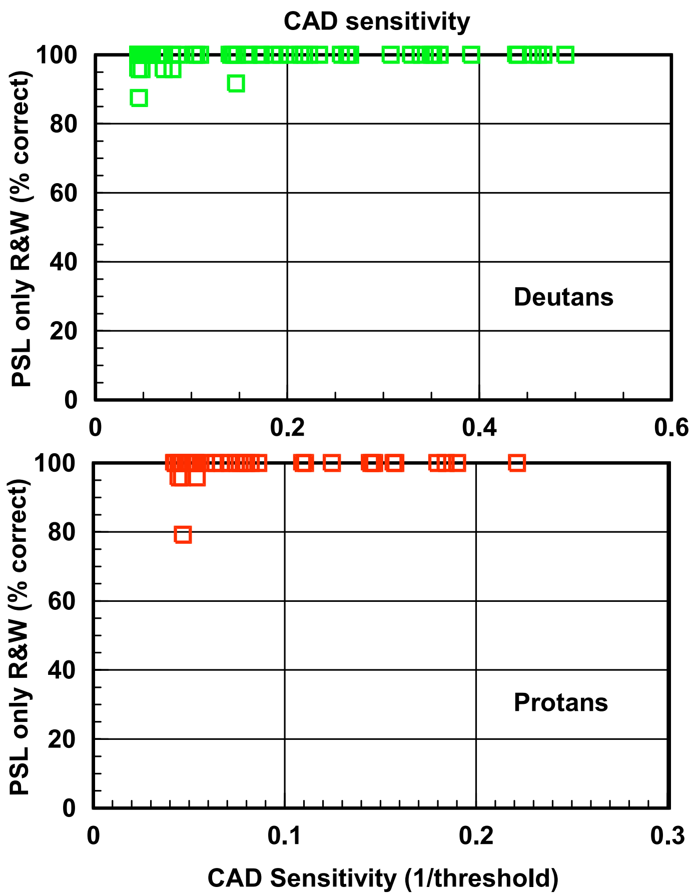

Figure 25: Graphs showing R=W and W=R errors only made on the PSL versus the RG CAD sensitivity.

defined as the reciprocal of the signal strength needed to just see the stimulus. The normalized CAD threshold falls within the range ~0.7 (i.e., better than normal sensitivity) to a maximum of 23.7 (a value limited by the phosphor limits of the display). Therefore, it is reasonable to use the reciprocal of the CAD threshold as a measure of chromatic sensitivity, which ranges from just above 1 for subjects with better than normal sensitivity, to ~0.04 for subjects with very slight or absent chromatic sensitivity. In the case of the most common occupational color vision tests, loss of chromatic sensitivity is more difficult to quantify because the tests do not measure the smallest color signal strength needed to detect the stimulus. Instead, most occupational tests yield scores of correct responses that are indicative of the subject's overall chromatic sensitivity. In spite of these limitations, it is of great interest to derive an index of mean chromatic sensitivity based on the subject's performance in several color vision tests. With this aim in mind, we used the parameters of each test to derive the best measure of average chromatic sensitivity, which ranges from around 1 for normal trichromats, to close to zero for subjects with very limited or complete absence of color discrimination:

Ishihara:	Fraction of plates named correctly
Dvorine:	Fraction of plates named correctly
ALT:	Fraction of stimulus pairs named correctly
Nagel:	See definition of RGI (i.e., red-green discrimination index)
CAD:	Reciprocal of threshold signal when measured in standard normal CAD units

Using this approach, we were able to quantify the subject's average chromatic sensitivity derived from the above tests in which we take the best available measure of the subject's overall ability to cope with a variety of color vision tasks. This measure of average chromatic sensitivity has been used to further justify the selection of minimum color vision requirements that can be classified as safe within the aviation environment and the exclusion of the very few subjects with poor overall RG chromatic sensitivity that pass the PAPI test.

Fig. 26 shows that the very few subjects that pass the PAPI with CAD sensitivities less than 0.2 (deutan) and 0.1 (protan) have poor overall chromatic sensitivity. The results also show that the pass/fail limits proposed on the basis of the CAD test ensure that the color deficient subjects that pass have an overall chromatic sensitivity greater than ~0.7 (deutans) and greater than ~0.5 (protans).

4.0 DISCUSSION

4.1 Color Vision Concerns in Aviation

Color enhances object conspicuity and can also be used to code information. In overcrowded displays or complex visual scenarios, color allows segmentation and grouping operations which enhance and speed up the acquisition and processing of visual information (Firth, 2001). The primary signal colors in aviation are red, green, and white, with blue and yellow as supplementary colors (CAA, 2006b). ICAO requires member nations to maintain a color standard to ensure pilots can recognize the colors of signal lights used in aviation (ICAO, 2006). Further, it acknowledges that the requirements are open to interpretation; and at present, there is no clearly defined line that specifies the class and severity of color deficiency beyond which the applicant is no longer safe to fly (for a full review, see CAA, 2006a). This is partly due to the varied task requirements and the technological advances in the use of color in the aviation industry.

There are a number of currently approved JAA and FAA color vision tests that applicants have to pass to be licensed to fly. In this investigation, we assessed 182 subjects using several color vision tests, and the results reveal both inter-subject variability and poor consistency

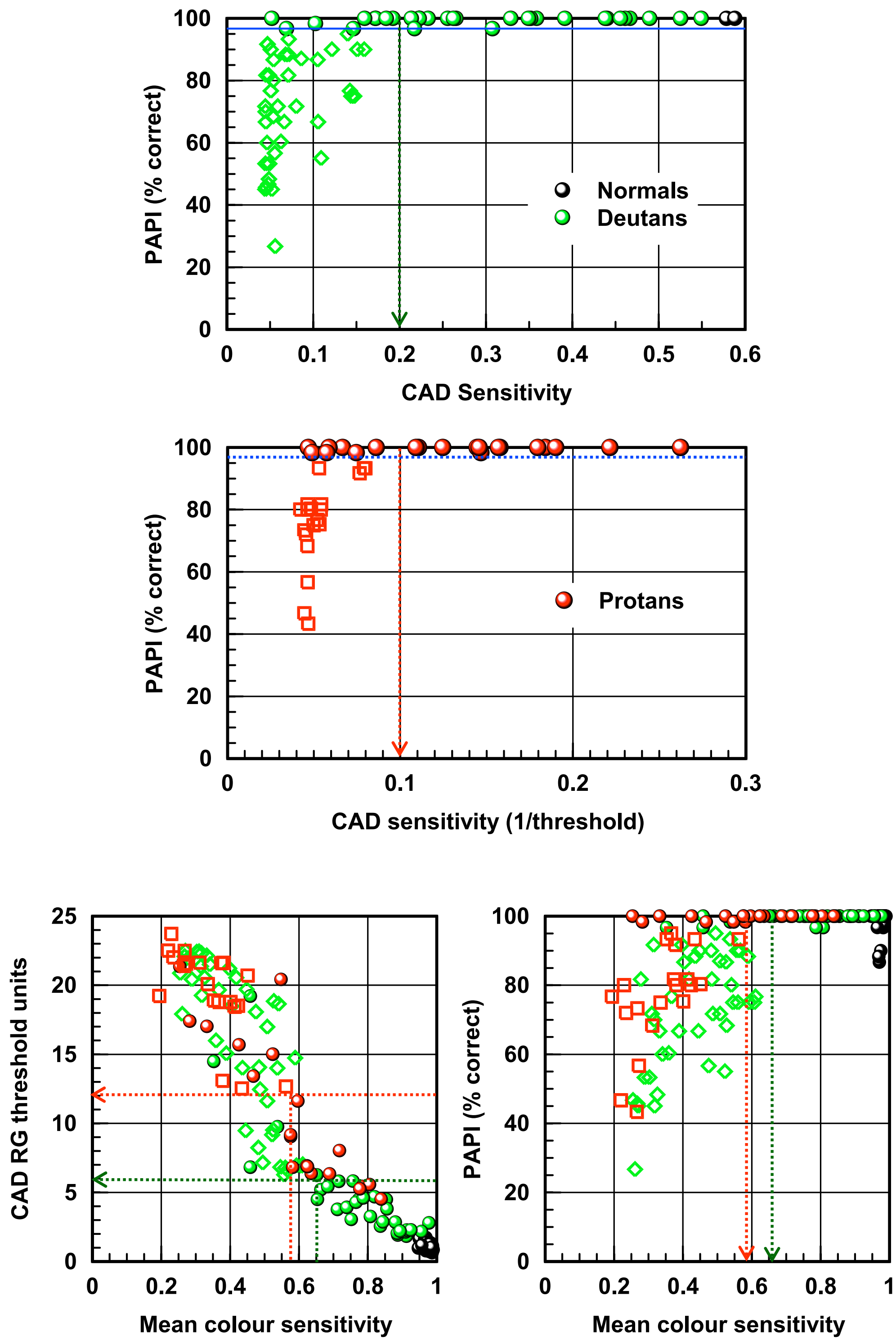

Figure 26: Justification for limits chosen. The top two diagrams compare PAPI (% correct) versus CAD test (sensitivity units=1/RG CAD threshold); color deficient subjects to the right of the dotted line perform the PAPI test as well as normal trichromats do. CAD sensitivity of 0.2 and 0.1 corresponds to 6 and 12 CAD threshold units, respectively, for deutans and protans. Note the change of horizontal scale units for protan observers. The bottom diagrams compare: left, CAD RG threshold units and right, PAPI performance, versus the "mean" or overall color sensitivity as computed by averaging the subject's performance on the Ishihara, Dvorine, ALT and Nagel anomaloscope tests.

among the various tests. Comparison of the outcome from Ishihara and the Dvorine tests shows that significant variability remains even when, at least in principle, the tests are very similar. These findings expose the facts that color deficient subjects can produce very different scores on the two tests and that normal trichromats can make more errors on one test, but not on the other. Further, the pseudoisochromatic tests designed for screening are not suitable to either accurately diagnose the class of deficiency, or to quantify the severity of color vision loss (Belcher et al., 1958; Birch, 1997).

The measures of chromatic discrimination on the CAD and Nagel anomaloscope reveal very poor correlation. Some color deficient observers that require a lot of red or green in the match can only accept a narrow, red-green range that is a characteristic of normal red/green color vision. Fig. 22 confirms well-known observations (Wright, 1946) in that the measure of chromatic sensitivity based on the Nagel range overestimates the subject's ability to discriminate red-green color differences under more natural conditions of illumination. Subjects with severe loss of chromatic sensitivity, as measured on the CAD test, can have RGI (Nagel) values that are similar to those measured in normal trichromats. The JAR pass criterion of 4 scale units (RGI=0.95) suggests that some normals are likely to fail, and of greater concern, that some severe color deficient subjects are likely to pass.

Fig. 21 shows comparisons between the ALT and the PAPI simulator tests. The results suggest that the ALT is a more challenging test, i.e. the average percent of correct scores are lower on the ALT than on the PAPI test. However, normals were found to make no errors on the ALT. The fact that color deficient subjects perform worse on the ALT than on the PAPI could be due to the increased number of colors presented on the ALT (three instead of two) and the very dark immediate surround in which the apertures of the lights presented are encased in the ALT.

These discrepancies among the various screening tests can be attributed partly to the differences in the methods of assessing color vision. These include differences in stimulus configuration, background lighting conditions, and also the different instructions given to subjects. The results from this study justify the recent concerns expressed by ICAO (1999) and CAA (2006a) in relation to some aspects of color assessment in aviation. The results confirm earlier findings from a study that examined three lantern types accepted by the JAA and found that the same individual can pass some of the tests and fail others (Squire et al., 2005). The study also showed 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 secondary color vision test is used, which varies between countries. The observed inconsistency in results between the currently approved color vision tests is therefore unsatisfactory, both in terms of flight safety issues, as well as being potentially unfair to some pilot applicants.

4.2 Advances in Assessment of Color Vision

Red-green chromatic sensitivity varies from "normal" performance to total "color-blindness", with an almost continuum of color impairment between these two extremes. Among congenital color deficient observers, the loss of RG color sensitivity varies along a continuous scale (Fig. 12). This is the most common type of color vision deficiency, affecting 8% of the male population (<1% females; see table in section 4.4). Yellow-blue deficiencies affecting the S-cone are very rare (1 in ~20,000) and are most often related to acquired color vision defects, as a result of eye disease, or as a side effect of toxicity or medication (see section 1.6 in this report). Acquired deficiencies tend to be age-dependent and when unnoticed, may compromise safety within certain occupations. Follow up color vision tests are carried out for some Class 1 renewal of medical certificates (JAR-FCL 3, 2000). Current civil aviation tests are not designed to detect or measure YB sensitivity, so any anomalies due to YB loss or disease are likely to remain undetected. Yellow-blue discrimination has also become more important because of the extensive use of different colors in the modern aviation environment.

4.3 Safe Color Vision Limits in Aviation

The PAPI system built for this investigation reproduces the red and white lights under conditions simulating the spectral composition, the angular subtense, and the retinal illuminance of real PAPI lights viewed from 5.54 km. Comparisons between the performance on the PAPI and CAD tests reveal the minimum level of chromatic discrimination, below which subjects with red/green color deficiency no longer perform the PAPI task with the same accuracy as normal trichromats. Fig. 23 shows that deutans and protans with a RG threshold of less than 6 and 12 CAD units, respectively, perform the PAPI simulator test in a similar way to normal trichromats. The PSL test addresses the question as to whether subjects can also name correctly the color of the PAPI lights when all lights are of the same color, either red or white. Fig. 25 shows how the ability to correctly name the white and red lights on the PSL test relates to the subject's RG discrimination sensitivity on the CAD test. Color deficient subjects do not have difficulty with this condition, and no subject with RG threshold units less than 6 and 12 for deutans and protans, respectively, confuse R=W or W=R when all four lights are the same color. The PAPI and PSL tests

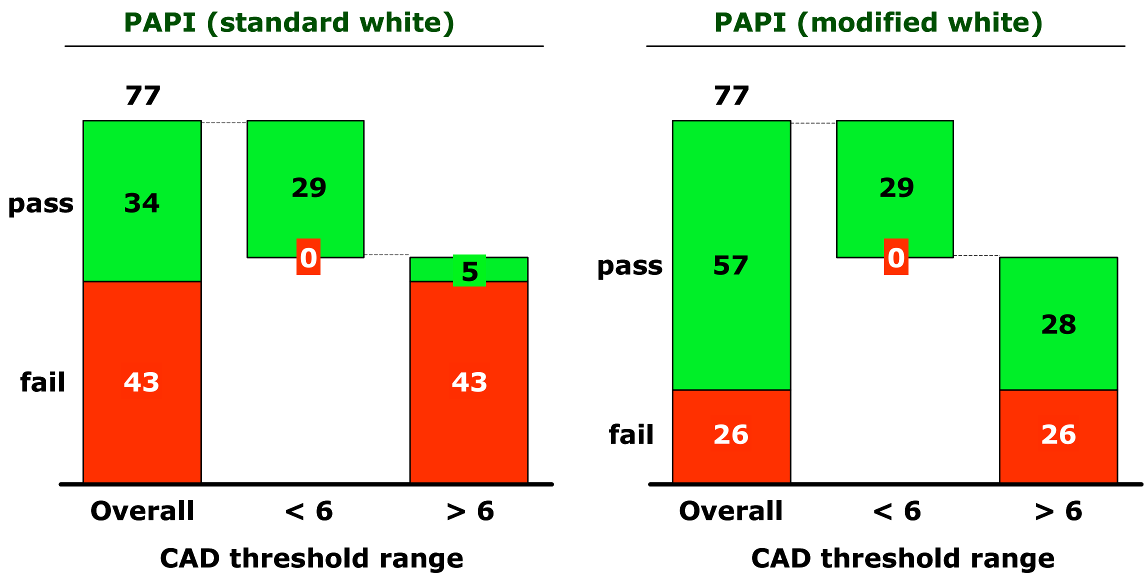

Figure 27: Summary of deutan subjects' results if the proposed pass/fail criterion of 6 RG CAD threshold units is accepted.

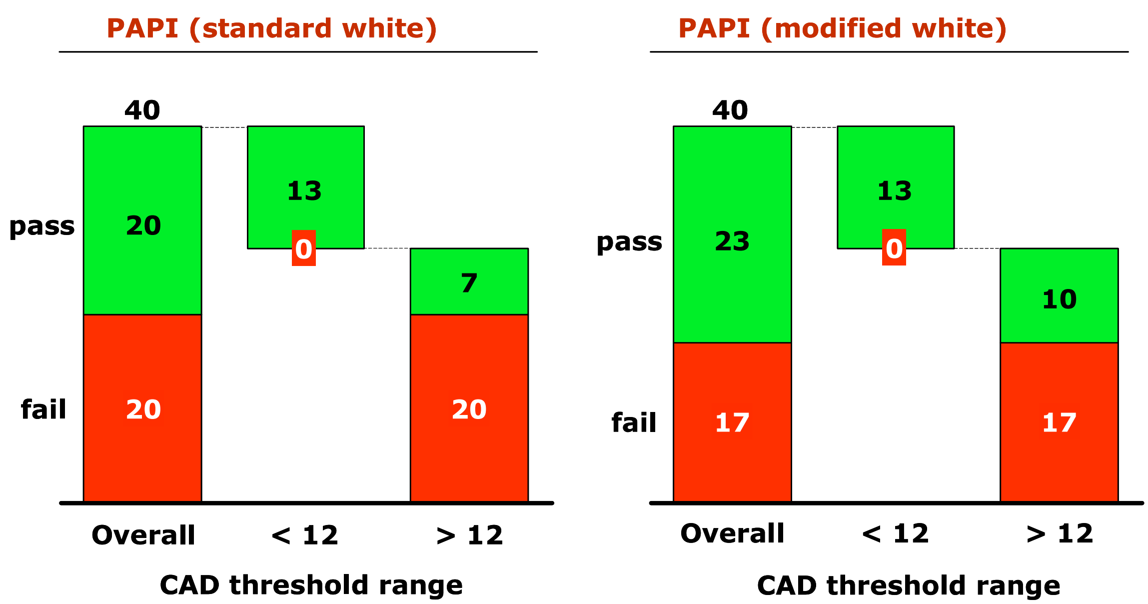

Figure 28: Summary of protan subjects' results if the proposed pass/fail criterion of 12 RG CAD threshold units is accepted.

also include a CC white to investigate whether the use of a more "bluish" white can improve performance. The results show that this is indeed the case for both PAPI and PSL tests in all subject groups.

Using the proposed pass and fail limits of 6 and 12 RG CAD threshold units for deutans and protans, respectively, 29 out of 77 deutans and 13 out of 40 protans would pass the PAPI simulator test using the standard white (Fig. 27). However, five deutans and seven protans (Fig. 28) with RG thresholds larger than the proposed safe limits also pass the PAPI. An important question is whether these subjects are disadvantaged unfairly if the new pass/fail limits were to be adopted. There is little doubt that these subjects have severely reduced RG color discrimination (as revealed in all the color vision tests employed in this investigation). The overall loss of chromatic sensitivity becomes increasingly more severe as the subject's thresholds increase beyond the recommended limits. These subjects are therefore likely to be disadvantaged in other, less safety-critical, visual tasks that involve color discrimination. By computing an average chromatic discrimination performance on a battery of color vision tests, we are able to examine whether these subjects have poor overall color discrimination performance. Fig. 26 shows that the very few subjects that pass the PAPI with CAD threshold limits greater than 6 (deutan) and 12 (protan) units have poor overall chromatic sensitivity. The results also show that the pass/fail limits proposed on the basis of the CAD test ensure that the color deficient subjects that pass have an overall RG chromatic sensitivity greater than ~0.7 (deutans) and greater than ~0.5 (protans).

The use of the color-corrected white condition increases quite significantly in the number of subjects that passed the PAPI test. For deutans (Fig. 27) with thresholds higher than 6 units, 28 subjects passed instead of only five with the standard white condition. Similarly, for protans (Fig. 28), ten passed in the modified condition, instead of seven for the standard white condition. Overall, the CC white condition improves performance of the PAPI test for all subject groups. This improvement is, however, more significant for deutan than protan color deficient observers.

The analysis shown graphically in Figs. 27 and 28 reveal that 37.6% of subjects with deutan-like deficiency and 32.5% of subjects with protan-like deficiency would be classified as safe to fly under the new research-based pass/fail limits. When using the current PAPI systems (i.e. PAPI with the standard white light), 44% of deutan subjects passed the PAPI test and, according to the proposed pass/fail limit of 6 SN CAD units, 85% of these subjects (i.e., the deutans that passed the PAPI) would be classified as safe to fly. Similarly, 50% of subjects with protan-like deficiency passed the PAPI test, and according to the pass/fail limit of 12 SN CAD units, 65% of these subjects would be classified as safe to fly.

So far, 386 color deficient observers have been examined on the CAD test as part of this study, as well as other ongoing studies. Two hundred fifty-five (255) of these subjects had deutan and the remaining 131 had protan deficiencies. If the proposed limits of 6 and 12 SN CAD units for deutan and protan subjects, respectively, were applied to this larger group, 36.1% of deutans and 29.8% of protans would pass these limits. These findings suggest that 35% of the total number of color deficient subjects investigated should be certified as safe to fly. The percentages shown above are very close to those estimated from the smaller number of color deficient subjects (n = 117) that participated in the PAPI study.

The proposed pass/fail limits relate accurate measurements of chromatic sensitivity to the subjects' performance on the most critical color vision task in aviation. The limits ensure that the subjects that pass can also perform the PAPI task with the same accuracy as normal trichromats. In addition, the proposed limits ensure that all subjects that fail according to these limits have poor, overall chromatic discrimination sensitivity.

4.4 Benefit Analysis of Using the New Approach

a. Analysis Based on CAA/JAA Pass Criteria and Guidelines

Table 1 shows the different classes and relative distribution of color deficient subjects that make up ~8% of the male population (Sharpe, et al. 1999).

The data in Table 2 are based on the strict CAA/JAA pass criteria for the Ishihara test. A similar analysis using the FAA pass/fail criteria for Dvorine, Ishihara, and ALT tests will be shown separately. Interestingly, only 7% of color deficients pass the Holmes-Wright (HW) lantern, and there is no clear evidence to suggest that all of these subjects are safe to fly.

One hundred sixty-three (163) color deficients (i.e., 110 deutan and 53 protan subjects) have been investigated using the HW lantern in all of our studies. All protans failed, but only 99 of 110 deutan subjects failed the HW lantern. These findings make it possible to predict the number of color deficient subjects that are classified as safe to fly using current assessment procedures.

Table 3 shows the predicted outcome when the same 1,000 applicants were examined on the CAD test and the pass/fail criteria employed are based on the findings

Table 1: Percentage of color deficient observers that fail Ishihara and HW tests.

Accepted Prevalence of Color Vision Deficiencies†						
Protanope	**Deuteranope**	**Tritanope**	**P-nomalous**	**D-nomalous**	**T-nomalous**	**Total**
1	1.1	0.002	1	4.9	0	8.002
†*Gegenfurtener, K.R. & Sharpe, L.T. "Color Vision, from Genes to Perception" : Cambridge University Press.*						
Other facts based on normal trichromats and colour deficient subjects studies at AVRC						
Percentage of normal trichromats that fail the Ishihara test*						15.8
Percentage of protan subjects that fail the Ishihara test						100
Percentage of deutan subjects that fail the Ishihara test						99
Percentage of normal trichromats that fail the Holmes-Wright lantern test						0
Percentage of protan subjects that fail the Holmes-Wright lantern test						100
Percentage of deutan subjects that fail the Holmes-Wright lantern test						90

*Results based on 202 normal trichromats when employing the strict CAA/JAA pass criteria (i.e., no errors, no misreadings in the first 15 plates of the 24-plate version).

Table 2: Predicted outcome per thousand applicants using current CAA/JAR guidelines.

Predicted outcome per 1000 applicants using current assessment methods				
Applicants	***1000***	***No. that fail Ishihara***	***No. that fail HW***	***No. classed as safe to fly***
Normals	920	145	0	920
Deutans	60	59	53	6
Protans	20	20	20	0
Total	**1000**	**225**	**73**	**926**
Percentage of applicants that undergo secondary tests =				**22**
Percentage of deutan subjects that pass secondary tests =				**10**
Percentage of protan subjects that pass secondary tests =				**0**
Percentage of total color deficient subjects that pass =				**7**

Table 3: Predicted outcome per thousand applicants using the new, CAD based pass/fail limits.

Predicted outcome per 1000 applicants using CAD pass / fail criteria**				
Applicants	***1000***	***No. that pass CAD as normals***	***No.that fail set CAD limits***	***No. classed as safe to fly***
Normals	920	920	0	920
Deutans	60	0	38	22
Protans	20	0	14	6
Total	**1000**	**920**	**52**	**948**
Percentage of applicants that undergo secondary tests =				**0**
Percentage of of deutan color deficients that pass =				**36**
Percentage of protan color deficients that pass =				**30**
Percentage of total color deficient subjects that pass =				**35**

** % deutan subjects that pass CAD (pass < 6 SNU) = 36.1
** % protan subjects that pass CAD (pass < 12 SNU) = 29.8

from this study. The current procedures that employ a secondary test whenever the applicant fails the primary test results in 22% of the applicants having to take the secondary test. The new procedure based on the CAD pass/fail limits does not require any secondary test. The analysis in Table 3 shows that 36% of deutan and 30% of protan subjects meet the pass/fail criteria established experimentally and can therefore be classified as safe to fly. Given the higher prevalence of deutan subjects within the male population, these findings suggest that 35% of all color deficient subjects pass the new guidelines and are therefore safe to fly.

If the HW secondary test is replaced by the Nagel anomaloscope test (using the current JAA pass/fail limits for midpoint and range), the percentage of color deficient subjects that are classified as safe to fly decreases to less than 2%. If the HW lantern test is replaced with other, less demanding secondary tests, the percentage of color deficient subjects that pass can increase to more than 10%. There is, however, no guarantee that the color deficient subjects that pass these secondary tests can carry out the most safety-critical, color-related tasks.

b. Analysis Based on FAA Pass/Fail Criteria and Guidelines

The FAA pass/fail criteria are more liberal and were selected deliberately to be fair to those color deficient applicants that may well be able to carry out safety-critical, color-related tasks with the same accuracy as normal trichromats. The criteria are not, however, based on the PAPI task and employ a large number of recommended tests that fail to correlate well when used to quantify the severity of color vision loss. This means that some of the applicants fail some tests and pass others. In this investigation, we have included two of FAA's most popular tests, the Ishihara and the Dvorine pseudoisochromatic plates. The ALT (i.e., a Farnsworth Lantern (FALANT) with modified filters designed to produce lights with chromaticities specified for aviation signal lights (Milburn & Mertens, 2004)) was included in the study. Normal trichromats pass both the FALANT and the ALT lanterns, but no data are available to describe how the deutan and protan subjects examined in this study perform on the FALANT test. The following analysis is, however, of great interest since it provides a useful comparison of expected outcomes when FAA guidelines are followed.

Table 4 lists the percentage of normals, deutan, and protan subjects that fail each of the three tests: Ishihara, Dvorine, and ALT, according to the FAA criteria. The first observation of interest is that the Dvorine test passes the largest number of deutan and protan subjects. The Ishihara is more demanding for both deutan and protan subjects, and all protan subjects fail the ALT test. In the following sections we examine the correlation between the outcome of each of these tests and the corresponding PAPI pass/fail scores for the 182 subjects investigated.

The contingency tables below (i.e., Table 5) show that only one deutan subject (i.e., less than 1%) passed the Ishihara according to the FAA criterion and failed the PAPI. Not unexpectedly, the same subject also passed the Dvorine test and failed the PAPI. Of equal interest is the observation that 31% of color deficients (deutan and protan) failed the Ishihara but passed the PAPI. The results also show that only 24% of the color deficients that failed the Dvorine test passed the PAPI. It is also of interest to investigate the outcome if the ALT were used as a secondary test for those deutan and protan pilot applicants that fail either Ishihara or Dvorine tests using the FAA pass/fail criteria.

This analysis should be strictly carried out using the applicant's performance data on the FALANT test, but in the absence of such data, the ALT test provides the

Table 4: Percentage of color deficient observers that fail Ishihara, Dvorine and ALT tests (using FAA pass/fail criteria for the Dvorine and the 24-plate Ishihara test).

Other percentages based on the FAA pass / fail guidelines*	
Percentage of normal trichromats that fail the Ishihara test	0
Percentage of protans that fail the Ishihara test	90
Percentage of deutans that fail the Ishihara test	81.8
Percentage of normal trichromats that fail the Dvorine plates	0
Percentage of protans that fail the Dvorine test	82.5
Percentage of deutans that fail the Dvorine test	75.3
Percentage of normal trichromats that fail the ALT	0
Percentage of protans that fail the ALT	100
Percentage of deutans that fail the ALT	77.9

*The results listed in the table are based on the FAA pass limits for Ishihara/Dvorine tests (fail equals 7 or more errors). Percentages based on 65 normal trichromats and 117 color deficients (i.e., 77 deutan and 40 protan subjects). All protans failed the ALT test, but only 60 out of 77 deutans failed the same test (pass classification on the ALT requires no more than one error on 27 trials).

Table 5: Contingency tables showing results of Ishihara, Dvorine and ALT tests and the corresponding pass/fail PAPI scores.

Number of subjects that pass/fail each test (i.e., a pass requires less than 7 errors on Ishihara; less than 7 errors on Dvorine; one or no errors on ALT)

Normals (65)		PAPI Pass	PAPI Fail	Protans (40)		PAPI P	PAPI F	Deutans (77)		PAPI P	PAPI F
Ishihara	Pass	62	3	Ishihara	P	4	0	Ishihara	P	13	1
	Fail	0	0		F	16	20		F	20	43

Normals (65)		PAPI P	PAPI F	Protans (40)		PAPI P	PAPI F	Deutans (77)		PAPI P	PAPI F
Dvorine	P	62	3	Dvorine	P	7	0	Dvorine	P	18	1
	F	0	0		F	13	20		F	15	43

Normals (65)		PAPI P	PAPI F	Protans (40)		PAPI P	PAPI F	Deutans (77)		PAPI P	PAPI F
ALT	P	62	3	ALT	P	0	0	ALT	P	17	0
	F	0	0		F	20	20		F	16	44

Table 6: Pass/fail scores on Ishihara and Dvorine compared against the ALT.

Number of subjects that pass/fail each test (i.e., a pass requires less than 7 errors on Ishihara; less than 7 errors on Dvorine; one or no errors on ALT)											
Normals (65)		ALT		**Protans (40)**		ALT		**Deutans (77)**		ALT	
		Pass	Fail			P	F			P	F
Ishihara	Pass	65	0	Ishihara	P	0	4	Ishihara	P	11	3
	Fail	0	0		F	0	36		F	6	57

Normals (65)		ALT		**Protans (40)**		ALT		**Deutans (77)**		ALT	
		P	F			P	F			P	F
Dvorine	P	65	0	Dvorine	P	0	7	Dvorine	P	13	6
	F	0	0		F	0	33		F	4	54

Table 7: Predicted outcome per thousand applicants when using Ishihara, Dvorine and ALT tests.

Predicted outcome per 1000 applicants using the FAA criteria							
Applicants	***1000***	***No. that fail Ishihara***	***No. that fail Dvorine***	***No. that fail ALT***	***Safe to fly Ishihara***	***Safe to fly Dvorine***	***Safe to fly ALT***
Normals	920	0	0	0	920	920	920
Deutans	60	49	45	47	11	15	13
Protans	20	18	17	20	2	3	0
Total	**1000**	**67**	**62**	**67**	**933**	**938**	**933**

Percentage of deutan subjects classed safe to fly according to each test	18.3	25.0	21.7
Percentage of protan subjects that are classed safe to fly	10.0	15.0	0.0
Percentage of all color deficient subjects classed safe to fly	**16.3**	**22.5**	**16.3**

nearest, useful substitute. Findings from a recent study (Cole & Maddocks, 2008) suggest that some subjects that pass the FALANT test can confuse red and white signals and may not therefore perform the PAPI task as well as normal trichromats. Table 6 compares the subjects pass/fail results on Ishihara and Dvorine against the corresponding pass/fail results on the ALT. The results show that it is possible to pass Ishihara or Dvorine and then to fail the ALT. In the case of Dvorine, six deutan and seven protan subjects pass the test only to fail the ALT. Interestingly, the reverse to this also happens. Four deutans that fail the Dvorine pass the ALT and six deutans that fail the Ishihara also pass the ALT.

These findings clearly reveal the lack of agreement that exists when using different occupational color screening tests and the difficulties involved if one wished to use such data to predict performance in other tasks, such as the PAPI.

Table 7 shows the predicted outcome per thousand applicants when using Ishihara, Dvorine, and ALT tests. Although the FALANT is an accepted FAA alternative test, the ALT test (used by FAA for assessing air traffic controllers) was employed in this study. The ALT fails all protan subjects and passes only 16% of deutan applicants. The results are of interest since they show clearly that an applicant is most likely to pass the Dvorine test, followed by ALT, and Ishihara when only deutan subjects are considered. In the case of protan deficiency, an applicant is most likely to pass the Dvorine, followed by Ishihara, but all protans fail the ALT test.

5.0 CONCLUSIONS

The aim of this project was to develop new methods for accurate assessment of color vision and to provide evidence-based guidelines for minimum color vision requirements for flight crew. The current diversity in color vision testing methods and standards demonstrates the need to adopt more objective assessment techniques internationally and to set minimum color vision requirements that are both safe and fair to the applicants.

The CAA/JAA guidelines are strict, and consequently only 10% of deutan applicants are classified as safe to fly (see Table 2). The applicants that pass are likely to perform the PAPI task as well as normal trichromats. The current CAA/JAA procedures have some disadvantages. A large percentage of applicants fail the primary tests and have to take lengthy secondary tests that differ in various member states. This introduces some variability in that the same subjects can pass some secondary tests and fail others. Some of these subjects may not necessarily be able to carry out the PAPI tasks as normal trichromats. Equally important, a large percentage of color deficients that can carry out the safety-critical, color-related tasks fail the tests and are, therefore, unfairly disadvantaged.

The FAA guidelines are more liberal, and as a result, more color deficient applicants pass. In the case of Dvorine and Ishihara tests, less than 1% of the applicants that pass these tests fail the PAPI. All color deficient applicants that pass the ALT test also pass the PAPI (see Table 5). The current FAA guidelines have some disadvantages. The color vision tests employed provide only an approximate measure of the severity of color vision loss, and the pass/fail limits have not been validated against the PAPI. The FAA accepts 14 different tests (FAA Guide for Aviation Medical Examiners, 2008). The use of so many tests that may not correlate well with each other (when used to assess the severity of color vision loss) increases variability and gives rise to odd outcomes when an applicant passes one test and fails another. Although the number of color deficient applicants that pass can increase significantly, this increase is test specific and often dependent on the use of the Signal Light Gun test that has not been validated against the PAPI task.

The results also show that one cannot regard subjects with deutan- and protan-like deficiencies as equivalent. Therefore, minimum color vision requirements must be set separately for each class of deficiency.

The principal findings of this study can be summarized as follows:

- Subjects with red/green congenital color deficiency can exhibit an almost continuous distribution of chromatic sensitivity loss.
- The loss of red/green chromatic sensitivity is greater in subjects with protan congenital color deficiency when compared to the deutan class. Unlike many conventional color vision tests, the CAD test cannot be learned, and hence the outcome depends entirely on the subject's chromatic sensitivity. The test provides the means to classify protan and deutan subjects and also quantifies the severity of color vision loss.
- Ninety-four percent of all applicants are likely to complete the CAD test in less than 20 seconds (using the fast-CAD option) which is based on screening for minimum color sensitivity in the deutan category. The remaining applicants are expected to fail fast-CAD. The definitive-CAD test takes 12 to 15 minutes to complete, and provides the information needed to establish whether the subjects that fail (i.e., ~4% of all applicants) have protan color deficiency and have residual chromatic sensitivity, within the established pass limit for protan subjects.
- Below ~60 years of age, normal aging does not significantly affect either RG or YB thresholds, provided adequate levels of ambient illumination are employed.
- Use of a modified "white" light results in significant, overall improvements in PAPI performance, particularly within normal trichromats and deuteranomalous observers.
- A comparison between the PAPI and CAD tests shows that deutan subjects with CAD thresholds <6 SN units and protan subjects with CAD thresholds <12 SN units can perform the PAPI test as well as normal trichromats.
- A small number of deutan and protan observers with thresholds higher than 6 and 12 SN units, respectively, passed the PAPI test, but these subjects exhibited poor overall chromatic sensitivity and are, therefore, likely to be affected unfavorably in other visual performance tasks that involve color discrimination.
- If these findings were adopted as pass/fail limits for pilots, ~35% of color deficient applicants would be certified as safe to fly.

6.0 REFERENCES

Alpern, M. (1979). Lack of uniformity in color matching. *J Physiol* 288, 85-105.

Alpern, M. & Pugh, E.N., Jr. (1977). Variation in the action spectrum of erythrolabe among deuteranopes. *J Physiol* 266, 613-46.

Barbur, J.L., Rodriguez-Carmona, M., Harlow, J.A., Mancuso, K., Neitz, J., & Neitz, M. (2008). A study of unusual Rayleigh matches in deutan deficiency. *Visual Neuroscience* 25, accepted for publication.

Barbur, J.L. (2003). Understanding color -normal and defective color vision. *Trends Cogn Sci* 7, 434-6.

Barbur, J.L. (2004). "Double-blindsight" revealed through the processing of color and luminance contrast defined motion signals. *Prog Brain Res* 144, 243-59.

Barbur, J.L., Cole, V.A., & Plant, G.T. (1997). Chromatic discrimination in subjects with both congenital and acquired color vision deficiencies.In *Color vision deficiencies XIII,* ed. Cavonius C. R., pp. 211-23. Kluver Academic Publishers, Dordrecht, Netherlands.

Barbur, J.L., Harlow, A.J., & Plant, G.T. (1994). Insights into the different exploits of color in the visual cortex. *Proc R Soc Lond B Biol Sci* 258, 327-34.

Barbur, J.L., Rodriguez-Carmona, M., & Harlow, A.J. (2006). Establishing the statistical limits of "Normal" chromatic sensitivity. *CIE Expert Symposium, CIE Proceedings 75 Years of the Standard Colorimetric Observer*.

Belcher, S.J., Greenshields, K.W., & Wright, W.D. (1958). Color vision survey using the Ishihara, Dvorine, Bostrom and Kugelberg, Bostrom, and American-Optical Hardy-Rand-Rittler tests. *Br J Ophthalmol* 42, 355-9.

Birch, J. (1997). Efficiency of the Ishihara test for identifying red-green color deficiency. *Ophthalmic Physiol Opt* 17, 403-8.

Carroll, J., Neitz, J., & Neitz, M. (2002). Estimates of L:M cone ratio from ERG flicker photometry and genetics. *J Vis* 2, 531-42.

CIE. Colors of Light Signals. Report S 004/E-2001. CIE Central Bureau, Vienna, Austria. 2001a.

CIE. International recommendations for color vision requirements for transport. Report 143. International Commission on Illumination, CIE technical report, Vienna, Austria. 2001b.

Civil Aviation Authority, Safety Regulation Group. CAP 168: Licensing of aerodromes, Chapter 6, 2004.

Civil Aviation Authority, Safety Regulation Group. Minimum color vision requirement for professional flight crew - Part 1: The use of color signals and the assessment of color vision requirements in aviation. Paper 2006/04. 2006a.

Civil Aviation Authority, Safety Regulation Group. Minimum color vision requirements for professional flight crew - Part 2: Task analysis. Paper 2006/04. 2006b.

Cole, B.L. (1993). Does defective colour vision really matter? In *Color vision deficiences XI*, ed. Drum, B., pp. 67-86. Kluwer Academic Publishers, Dordrecht, Netherlands.

Cole, B.L. (2004). The handicap of abnormal color vision. *Clin Exp Optom* 87, 258-75.

Cole, B.L. & Maddocks, J.D. (1995). Protans and PAPI: Recognition of a two color code by persons with defective color vision. In *Color vision deficiencies XII*, ed. Drum, B., pp. 501-10. Kluwer Academic Publishers, Dordrecht, Netherlands.

Cole B.L. & Maddocks J.D. (2008). Color vision testing by Farnsworth lantern and ability to identify approach-path signal colors. *Aviat Space Environ Med*, 79, 585-90.

Federal Aviation Administration, Version V of the Guide for aviation medical examiners (2008). Examination techniques and criteria for qualification, Item 52: Color Vision. www.faa.gov/about/office_org/headquarters_offices/avs/offices/aam/ame/guide/app_process/exam_tech/item52/amd/.

Firth, J. Color vision concerns in aviation, clinical concerns on color coding in aviation. NATO, RTO Technical Report 16. Operational Color Vision in Modern Aviation Environment. 2001.

Frey, RG. (1958). Most suitable pseudoisochromatic table for practice. Comparative studies on applicability of Bostrom, Bostrom-Kugelber, Hardy-Rand-Ritter, Ishihara, Rabkin, and Stilling tables. *Albrecht Von Graefes Arch.Ophthalmol* 160, 301-20.

Heath, G. & Schmidt, I. (1959). Signal color recognition by color defective observers. *Am J Optom Arch Am Acad Optom.* 36, 421-7.

ICAO. Review of the visual and color perception requirements for medical certification. ANC Task No. MED-9601. 11-1-1999.

International Civil Aviation Organization. Technical airworthiness manual. (2nd ed.-1987 Document 9051-AN/896), Amendment 10. Montreal, Quebec, Canada. 1988.

International Civil Aviation Organization (ICAO). International Standards and Recommended Practices. Annex 1 to the convention on International Civil Aviation - Personnel Licensing. 10th ed. 2006.

JAR-FCL 3. Flight Crew Licensing (Medical). Section 2 Subparts A, B, and C - JAA Manual of Civil Aviation Medicine. Chapter 17. 2000.

JAR-FCL 3. Flight Crew Licensing (Medical). Section 1 Subpart B - Class 1 Medical Requirements. (Amendment 2). 2002.

Milburn, N.J. & Mertens, H.W. Predictive validity of the Aviation Lights Test for testing pilots with color vision deficiencies. DOT/FAA/AM-04/14. 2004.

National Research Council-National Academy of Sciences, Committee on Vision. Procedures for testing color vision. Report of working group 41. Washington, DC. 1981.

National Transportation Safety Board. FedEx Boeing 727 crashed during landing at Tallahassee, Florida. Safety recommendation: NTSB/AAR-04/02. 2004.

Neitz, J. & Jacobs, G.H. (1986). Polymorphism of the long-wavelength cone in normal human color vision. *Nature* 323, 623-5.

Rodriguez-Carmona, M. (2006). *Variability of chromatic sensitivity: Fundamental studies and clinical applications.* PhD thesis, City University, London, United Kingdom.

Rodriguez-Carmona, M., Harlow, A.J., Walker, G., & Barbur, J.L. (2005). The variability of normal trichromatic vision and the establishment of the "normal" range. *Proceedings of 10th Congress of the International Color Association, Granada (Granada, 2005)* 979-82.

Sharpe, L.T., Stockman, A., Jagle, H., & Nathans, J. (1999). Opsin genes, cone photopigments, color vision, and color blindness. In *Color vision: From genes to perception*, eds. Gegenfurtner, K.R. & Sharpe, L.T., pp. 3-52. Cambridge University Press, Cambridge.

Sloan, L.L. & Habel, A. (1956). Tests for color deficiency based on the pseudoisochromatic principle; a comparative study of several new tests. *AMA Arch Ophthalmol* 55, 229-39.

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, 421-9.

U.S. Department of Transportation - Federal Aviation Administration. Federal Aviation Regulations. Part 23, Sec 23.1397. Washington, DC. 1988.

Vingrys, A.J. & Cole, B.L. (1986). Origins of color vision standards within the transport industry. *Ophthalmic Physiol Opt* 6, 369-75.

Vingrys, A.J. & Cole, B.L. (1993). The ability of color vision defective observers to recognize an optimised set of red, green, and white signals lights. In *Color vision deficiencies XI*, ed. Drum, B., pp. 87-95. Kluwer Academic Publishers, Dordrecht, Netherlands.

Wright, W.D. (1946). *Researches on normal and defective color vision,* Henry Kimpton, London.