Medical

Analysis of European colour vision certification requirements for air traffic control officers

CAP 1429

Published by the Civil Aviation Authority, 2016

Civil Aviation Authority,

Aviation House,

Gatwick Airport South,

West Sussex,

RH6 0YR.

First published 2016

Enquiries regarding the content of this publication should be addressed to: Sally Evans, medicalweb@caa.co.uk

Compiled by:

Professor John Barbur, Dr Marisa Rodriguez-Carmona and Joseph Hickey

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

Dr Sally Evans and Dr Adrian Chorley

UK Civil Aviation Authority, Safety and Airspace Regulation Group, Gatwick Airport South, RH6 0YR.

The latest version of this document is available in electronic format at www.caa.co.uk, where you may also register for e-mail notification of amendments.

Contents

List of figures and tables

Abbreviations

ATC	Air Traffic Controller
AVRC	Applied Vision Research Centre (City University)
AVOT	Advanced Vision and Optometric Tests
CAA	Civil Aviation Authority
CAD	Colour Assessment and Diagnosis test
CIE	Commission Internationale de l’Eclairage
CRATO	Colour Requirements for Air Traffic Operators test
CS	Chromatic Sensitivity
FAA	Federal Aviation Administration
ICAO	International Civil Aviation Organisation
JAA	Joint Aviation Authorities
JAR	Joint Aviation Requirements
L-cones	Long-wavelength sensitive cones
LC	Luminance Contrast
M-cones	Medium-wavelength sensitive cones
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

For a full list of relevant definitions, refer to CAA Paper 2006/04 (Civil Aviation Authority (CAA) 2006).

Nomenclature

°	degrees
cd m^{-2}	candelas per square metre
λ	wavelength (lambda), nm
λ_{max}	maximum (peak) wavelength of V(λ)
%	percent
2’ arc	2 minutes of arc
A	ampere (amp) unit of electric current
km	Kilometre
mm	millimetres (1 mm=10^{-3} of a metre)
nm	nanometres (1 nm=10^{-9} of a metre)
s	second (time)
μ	micro = x10^{-6}
V(λ)	Standard photopic luminous efficiency (for high ambient illumination) (CIE, 1924)
V’(λ)	Standard scotopic luminous efficiency (when very low light levels are involved) (CIE, 1951)

Executive summary

This report describes the findings of a project sponsored by the United Kingdom Civil Aviation Authority (UK CAA), the Colt Foundation and City University London.

Background

The use of colour in Air Traffic Control (ATC) applications has increased greatly during the past two decades largely as a result of advances in colour display technologies and lighting systems. There are numerous advantages to using colour in complex visual display applications, particularly when large visual fields are involved, as well as in lighting and signalling of information such as in airport lighting systems. The aim of this study was to examine and quantify the advantages of using colour with emphasis on ATC applications. In addition, we also aimed to quantify the relationship between severity of colour vision loss and the corresponding changes in visual performance when colour signals are involved.

Colour is arguably a very effective and compelling, but also attractive and efficient method available for enhancing visual performance when using visual displays. The appropriate use of colour signals can speed up visual search and hence reduce the time needed to locate objects and to absorb information in crowded scenes. Colour coding has also been shown to be superior to other achromatic visual attributes in many tasks that involve the processing of visual information over a large visual field, particularly when ‘grouping’ operations are involved in complex scenes. A fundamental property of colour mechanisms in human vision that yields significant advantages in visual search is the independent processing of colour and luminance contrast signals. The existence of different visual mechanisms dedicated to the processing of colour signals also means that objects defined by luminance and colour are resilient to background clutter and are often picked up instantly and processed in parallel over large regions of the visual field. There is little doubt that the use of colour signals can enhance visual performance and subsequent motor responses in visually demanding tasks. In spite of significant progress, the need to quantify the advantages of using colour signals and to establish safe and fair minimum colour vision requirements within visually demanding occupations continue to remain important challenges in human factors research.

The aim of this project was to examine and quantify the advantages of using both red / green (RG) and yellow / blue (YB) colour signals in large field, visual displays in subjects with normal trichromatic colour vision and to establish how these advantages change in subjects with measured levels of congenital loss of RG chromatic sensitivity. The Colour Assessment and Diagnosis (CAD) test has been used throughout to quantify each subject's RG and YB chromatic sensitivity.

New findings

The enhancement of key aspects of visual performance through the use of RG and / or YB colour signals has been investigated in large field visual displays involving visual search tasks. These include 'pop-out' in visual search, enhancement of target conspicuity, signalling information by means of colour and grouping / segmentation of spatial content into categories by colour. Task Completion Times (TCTs) and the subject's corresponding Correct Performance Scores (CPS) have been measured for targets defined by luminance contrast with or without the addition of RG and / or YB colour signals. The results confirm well established findings which show the importance of luminance contrast and the use of colour signals in visual search. In addition, the following new experimental findings have emerged from the study:

- Targets with higher luminance contrast can be detected quicker and easier than those of lower contrast, but the addition of colour signals to such targets can greatly reduce TCTs and also improve task performance accuracy. This is particularly obvious when the visual tasks involve the use of large visual fields and suprathreshold coloured targets and chromatic saturations typical of those employed in current ATC applications.

- In normal trichromats, TCTs decrease gradually with increasing colour signal strength with little additional benefit above 10 to 12 standard normal (SN) CAD threshold units. This is also the case when the target is defined by spatial cues and the task can be carried out (although more slowly) in the absence of colour signals.

- Both RG and YB colour signals yield significant advantages by shortening TCTs (often as much as four fold), even when colour is used redundantly and the task can be completed without the use of colour signals.
- When task specific information is displayed over large visual fields, YB colour signals have some advantage over RG, largely because YB chromatic sensitivity falls off less rapidly with increasing distance on the retina (measured as the angular separation between the location of the target and the point of regard). Although RG signals, particularly when small targets are involved, have advantages over YB signals in central vision, the opposite seems to be the case when the working visual field is large.
- Colour signals are more effective when added to targets defined by increments in luminance (i.e., when viewing bright as opposed to dark objects presented against a uniform background). This observation applies to both RG and YB stimuli.
- The loss of colour vision was assessed accurately in subjects with varying severity of congenital colour deficiency together with the corresponding loss in visual performance. In general subjects with even mild congenital RG colour deficiency perform less well when the task involves the use of colours of low chromatic saturation which they confuse, i.e., colours that differ mostly in RG content. The same mildly deficient subjects perform the same task as well as normal trichromats when YB colour signals are employed.
- Subjects with mild congenital colour deficiency (e.g. those with thresholds less than ~ 4 SN CAD units) can perform colour related tasks when several coloured targets are involved, but only when larger chromatic saturations are employed (i.e., > 10 SN CAD units). The addition of YB colour difference signals to targets defined by luminance and RG colour contrast ensures that mild congenital colour deficients perform visual

search tasks with virtually the same speed and accuracy as normal trichromats. Under such conditions, subjects with mild RG congenital deficiencies perceive suprathreshold colours defined by a normal YB component and a reduced RG signal.

- Visual performance in dichromats and also in subjects with severe loss of RG colour vision is significantly worse when compared to normal trichromats except for colours that rely almost entirely on YB colour differences. The ‘effective’ luminance contrast of red and green coloured targets in subjects with severe RG colour deficiency also varies with increasing chromatic saturation, even when the luminance contrast, as seen by normal trichromats, remains unchanged. This can have unexpected results on visual performance since increasing the chromatic saturation towards red or green will have no effect on RG colour contrast in dichromats, but such changes can alter significantly the luminance contrast of the target (as seen by the colour deficient subject). This effect can either cause an improvement or worsening of visual performance in an anomalous trichromat, although the effects will be largest in dichromats. The type of outcome depends on the size and polarity of the initial luminance contrast (as specified for a normal trichromat), type of colour deficiency involved (i.e., deutan or protan) and whether the increased colour saturation is in the red or green direction.
- By studying a large number of deutan- and protan-like subjects with varying severity of colour vision loss we were able to derive some equivalence between the subject’s performance on Holmes-Wright type A (HW-A) lantern and the corresponding CAD threshold. In addition, analysis of historical evidence of certification outcomes based on conventional tests (i.e., Ishihara pseudoisochromatic plates and the HW-A lantern tests) was also carried out. This analysis and the new results led to the establishment of the statistical equivalence between historically accepted pass / fail limits (based on the use of conventional tests) and the severity of colour vision loss as measured on the CAD test. We were therefore able to propose CAD threshold limits that correspond to 100% correct performance on the HW-A lantern. This approach yields justifiable

'colour safe' standards based on CAD thresholds without the need for detailed studies.

These new findings and the analysis of results obtained in normal trichromats and in subjects with congenital colour deficiency led to three categories that one can justifiably use to describe trichromatic colour vision:

1. **'Normal' trichromatic Colour Vision (CV1)**. This category includes all those with RG and YB CAD thresholds below the upper normal limits that have been established for healthy aging (Barbur and Rodriguez-Carmona 2015).

2. '**Functionally normal' trichromatic colour vision (CV2)**. This category includes all applicants with a CAD threshold ≤ 2.35 CAD units. This limit is sufficient to pass all normal trichromats, irrespective of age and ~ 7% of the least affected deutans. The latter exhibit almost normal RG colour discrimination and **pass** the HW-A lantern test **with zero errors**. In terms of anomaloscope match parameters, the deutans that pass exhibit match ranges within normal limits, but require more 'green' in the red / green mixture field to match the monochromatic yellow field. These subjects are not likely to have any colour detection and discrimination problems when suprathreshold colours defined by both RG and YB components are employed in visual displays. The least affected protan-like subjects make errors on the HW-A lantern and exhibit minimum RG colour thresholds well above 2.35 CAD units. No protan subjects can therefore be included in this category.

3. **'Safe' trichromatic colour vision (CV3)**. This category includes all applicants with YB CAD thresholds within the normal range and RG thresholds ≤ 4 CAD units. The higher limit is sufficient to pass all normal trichromats and ~ 22% of deutan subjects. This higher limit matches the percentage of deutans that pass the HW-A lantern (22%) when using the CIE protocol that has been recommended for use with this lantern. Although some of the deutans included in this group will have RG colour discrimination difficulties with small RG colour signals that are close to normal thresholds, all these subjects exhibit normal levels of visual

performance when suprathreshold colours defined by both RG and YB components are employed in visual displays. The least affected protan-like subjects make errors on the HW-A lantern and exhibit minimum RG colour thresholds above 4 CAD units. As a result, no protan subjects would be included.

Important recommendations

The enhancement of key aspects of visual performance as a result of adding RG and / or YB colour signals to targets defined by luminance contrast in large field visual tasks has been investigated in normal trichromats and in subjects with congenital colour deficiency. Speed of performance and accuracy have been measured and related to the subject's RG and YB colour vision sensitivity. The recommendations put forward in this report are based largely on these findings and also on the correlation between CAD thresholds and pass / fail error scores on Ishihara test plates and HW-A lantern tests measured in over 1000 subjects with both normal trichromacy and congenital colour deficiency.

In the absence of detailed studies designed to establish minimum colour vision requirements for specific occupational tasks (as has been done for flight crew in aviation and for train drivers for Transport for London), an acceptable alternative would be to consider carefully the three categories described above and to select the one that can be considered safe, without discriminating unfairly against those subjects with congenital colour deficiencies that can achieve levels of performance equivalent to normal trichromats.

- If the visual task requires detection and naming of colours for small signal lights (e.g. red, green, yellow, blue and white, etc.), or the discrimination of the smallest possible colour differences in order to judge uniformity of colour reproduction in manufactured goods, or the need to adhere to the commonest appreciation of perceived colour appearance and colour names and / or the ability to use efficiently faint, desaturated colours to segment objects into groups on visual displays, a **CV1** pass would be justified.

- When the visually-demanding, colour-related tasks involve the use of suprathreshold colours (e.g. with chromatic saturations and colour differences that are well above detection thresholds, which is the normal practice when colour is used deliberately to enhance visual performance in working environments), a **CV2** pass can be allowed without compromising either efficiency or safety. Applicants with a CV2 pass can discriminate small RG colour differences, make few errors on the Ishihara plates test, and more importantly, make no errors with signal lights that are equivalent to the HW-A lantern (i.e., small (often diffraction limited) whites, greens and reds of varying intensity). Based on these findings, a CV2 pass would be appropriate for air traffic controllers and seafarers (i.e., lookout officers).
- The **CV3** category is appropriate for the majority of normal working environments that employ suprathreshold colours and do not involve discrimination of fine colour differences or the need to make correct judgements of colour appearance. Subjects with a CV3 pass will have sufficient RG chromatic sensitivity to carry out suprathreshold, colour-related tasks, even when RG colour signals align along the corresponding colour confusion lines. CV3 is an important category for a number of reasons and hence benefits from further justification. Visual scenes in real working environments involve the use of large objects with spatial features that are often several times above the acuity limit. This is also the case when objects and images are generated on visual displays. The overall appearance of each object is determined by a combination of its size, luminance and YB and RG chromatic contrast. Differences in YB colour signals are often present in objects that are commonly classed as reddish and greenish and in red / white signal colours (such as the Precision Approach Pathway Indicator (PAPI) lights). If large chromatic saturations are employed, subjects with mild RG colour deficiency (e.g. those with a CV3 pass) will be able to make use of the reduced RG colour signal to carry out the colour-related task, even in those rare cases when YB colour differences are absent. The only slight disadvantage is that these subjects will be slower than normal trichromats when the tasks require visual

search in large displays. One can, however, make use of this information and the novel findings that have emerged from this study in relation to the increased efficiency of YB signals in the periphery of the visual field to optimise the colours used in visual displays. When suprathreshold YB colour difference signals are also added to objects defined by luminance and RG colour contrast, congenital deficients that fall into the CV3 category can perform multi-colour visual search tasks with the same accuracy and speed as normal trichromats. Equally importantly, the majority of these subjects pass the HW-A lantern test, with equal number of false positives and false negatives centred with respect to the pass / fail limit of 4 CAD SN units. A CV3 pass is therefore appropriate for applications that involve the use of large colour differences, particularly when these also involve YB colour signals. With appropriate design and choice of colours, the CV3 category can also be appropriate for use in the ATC environment as well as in many other occupations that involve the use visual displays.

Chapter 1

Introduction

Air traffic control officers (ATCOs) are tasked with providing the aeronautical guidance needed for safe and efficient movement of air traffic from origin to destination. The primary purpose of air traffic control (ATC) worldwide is to provide information and support for pilots to ensure safe transit of aircraft and to organise and expedite the flow of traffic.

The immediate airport environment is controlled by visual observation from the airport control tower, including movement of aircraft and vehicles operating on taxiways and runways of the airport itself and aircraft in the air near the airport (anywhere from 10 to 20 km). Surveillance displays are also available to controllers for airborne traffic approaching and departing. These dynamic displays include a map of the area, the position of various aircraft, and data tags with aircraft identification, speed, altitude and other relevant information. En-route air traffic controllers work in facilities called ATC Centres and provide services to aircraft in flight between airports. Each centre is responsible for many thousands of square kilometres of airspace, known as a Flight Information Region, and for the airports within that airspace. The National Air Traffic Services, commonly referred to as NATS, is the main air navigation service provider in the United Kingdom. It provides en-route air traffic control services to flights within the UK Flight Information Regions and the Swanwick Oceanic Control Area (north eastern part of the Atlantic Ocean). NATS also provides ATC services to fourteen UK airports.

The main problems faced by ATC services are primarily related to peak volumes of air traffic demand and weather changes. Several factors dictate the amount of traffic that can land at an airport in a given period of time. About four minutes is allowed for each landing aircraft to touch down, slow, and exit the runway before the next aircraft crosses the approach end of the runway. Further challenges can also arise when departures are slotted in the time between arrivals. Advances in ATC systems have now made it possible to sequence planes hours in advance; nonetheless it is critical to monitor cruising altitude and horizontal separation between airborne aircraft.

Weather is also a major factor in traffic capacity. Rain, ice or snow on the runway cause landing aircraft to take longer to slow down and exit, thus reducing the safe arrival rate and requiring more space between landing aircraft. Fog can also cause a decrease in landing rate and thunderstorms present a variety of hazards to aircraft and disruption to prearranged sequences. As a result, there is an unavoidable need for ATCOs to handle a large amount of information on visual displays and to sustain attention to visually-demanding operations over long periods of time.

This report follows on from a previous CAA report on “Minimum Colour Vision Requirements for Professional Flight Crew: II. The Use of Colour Signals and the Assessment of Colour Vision Requirements in Aviation” published in 2009. Although this mainly involved professional pilots many of the concepts and methodology employed are similar.

1.1 The use of colour in ATC applications

Colour is arguably a very effective, compelling, and attractive method available for enhancing visual performance when using visual displays. The appropriate use of colour signals can speed up visual search and hence reduce the time needed to locate objects in crowded scenes (Carter 1982). Evidence that has emerged from several studies demonstrates clearly that colour coding is superior to other achromatic visual attributes in many tasks that involve the processing of visual information over a large visual field, particularly when ‘grouping’ operations are involved in complex visual scenes (Christ 1975). A fundamental property of colour mechanisms in human vision that yields significant advantages in visual search is the independent processing of colour and luminance contrast signals (Barbur et al. 1994; Birch et al. 1992; Kaiser and Boynton 1996). The existence of different visual mechanisms dedicated to the processing of colour signals also means that objects defined by colour are resilient to background clutter and are often picked up instantly and processed in parallel over large regions of the visual field (Barbur et al. 2003; Barbur and Forsyth 1988; Treisman and Gelade 1980). Colour signals can therefore be used as a distinct, additional feature to increase target ‘conspicuity’, but also to signal information and to group together spatially discrete objects that are usually defined by luminance contrast. Although the mechanisms that process luminance contrast have a much higher spatial resolution and enable us to see fine edges and contours and to detect very small spatial detail in low contrast, the addition of colour signals can enhance the conspicuity of objects of low luminance contrast (Barbur and Forsyth 1988), but more importantly, different colours can be used to signal information through colour coding and also to segment objects into groups and the visual field into areas of interest. All these advantageous attributes can enhance visual performance and speed up motor responses in visually demanding tasks.

In the past, the primary colour-related task of ATCOs involved the identification of coloured text on flight progress strips and recognition of aircraft and their direction of flight at night. The latter was performed from the ATC tower and was based largely on perception of red, white and green navigation lights. These early techniques were important since the use of colour was not usually accompanied by redundant cues (Mertens 1990). The use of colour has increased greatly in the ATC environment and has become an important tool in arranging and presenting information as a result of rapid advancement in applications and the use of visual displays. New technologies and automation tools have been added to existing displays which allow users to customise their own colour schemes. Consequently, many recommendations as to the use of colour for ATC followed (HF-STD-001 2003; CAA 2014; Cardosi and Hannon 1999; HF-STD-002 2007). In general, ATCOs have to process a large amount of information and even when used redundantly, colour signals can help speed up visual processing, enhance performance and enable operators to cope with more challenging tasks. Knowledge on the use of colour in visual displays has increased significantly during the last few decades and largely pragmatic guidelines have emerged to maximise the benefits of colour coding. The following is a summary of some known guidelines:

- Whenever colour is used on a display for coding safety critical information, it should be used redundantly with additional cues, i.e. spatially unique features defined by luminance contrast, moving or briefly presented, repeated flashes, audible signals, etc.
- Colour should be used consistently for the same functions across all the displays used by a single controller. The same colour conventions and meanings assigned to individual colours also need to be compatible across displays.

- Historical / cultural colour conventions should not be violated, such as red for danger, yellow for warning and green for clear.
- When a specific colour is used to assign a unique meaning, the number of unique colours employed should be less than six.
- All colour coded text and symbols should be at least three times above the normal acuity limit and presented in sufficient luminance contrast.
- The specific colours and background luminance selected for use on a visual display must also take into account the environment, ambient lighting and the limits of the specific monitors.
- Pure, bright, highly saturated colours should be used sparingly. These colours should only be used for displaying critical information briefly, so as to avoid perceptual habituation and strong chromatic adaptation. Saturated red and blue colours, when presented simultaneously, can create unwanted issues with depth perception.
- Saturated blue should not be used for displaying text or other small symbols and fine spatial details and blue should in general not be used as a background colour.

Although existing guidelines are undoubtedly useful, they also have limitations since they are largely concerned with the perception of colour and are not always based on studies designed to evaluate the effect of colour on task performance (Xing and Schroeder 2006). Enhancement of sustained attention, parallel processing of colour defined features with immediate identification of coloured objects, signalling of specific information by means of colour coding and grouping operations by segmenting objects of interest into groups and / or useful categories are important attributes that benefit from appropriate use of colour signals. Other benefits of colour vision in relation to large field, visually demanding tasks have also been identified. The work carried out at the Centre for Applied Vision Research at City University on the use of colour in displays also identified the advantages of enhanced conspicuity that can be achieved by adding colour signals to objects defined by luminance contrast (Barbur and Forsyth 1988; Walkey et al. 2005), the relative benefits of using RG and YB signals when the visual scene involves the use of large visual fields and the contribution colour signals can make to reaction times (Barbur et al. 1998; Walkey et al. 2006). In this study we have identified the most important benefits of using colour in visual displays (with emphasis on ATC applications) and have developed visual tests to quantify these benefits in subjects with normal trichromatic vision. In general, colour signals were used redundantly so that completion of the task was possible even in the absence of colour. In addition, we also investigated a large number of subjects with congenital colour deficiency and examined how the severity of colour vision loss affects the levels of performance measured in normal trichromats. The following sections describe in greater detail the roles colour signals can play in complex visual tasks.

Enhancement of object conspicuity

An important property of colour signals is to enhance the ‘effective’ contrast of objects. An object defined by both luminance and chromatic contrast, in general has a higher ‘perceived’ contrast. In the absence of colour signals, visual search times and other aspects of visual performance that involve detection and recognition of spatial cues depend largely on the luminance contrast of the target. When the luminance contrast is low, the addition of colour signals, particularly to targets defined by luminance increments, results in improved visual performance and shorter task completion times (Barbur and Forsyth 1988; Walkey et al. 2005).

Pop-out and parallel processing of colour signals

ATCOs are often required to detect and interpret quickly novel, critical information that may appear in different sections of an ATC display. This is usually not a problem on a simple, un-crowded display, but when the display is cluttered with additional information, visual crowding can make the visual search more challenging. Crowding caused by objects and characters defined by luminance contrast has less effect on coloured stimuli which can still be detected and localised rapidly. This phenomenon is called “pop-out” and has been described in earlier studies (Treisman and Gelade 1980). Pop-out is especially useful when crowded and large visual displays are employed. Many stimulus features that enable spatial judgements such as the presence of a gap in a ring can only be detected in central vision within a small visual field around the point of regard, often described as ‘visual lobe size’. Eye-movements shift the point of regard over the scene until the target of interest is detected. When the visual lobe size is very small (as is the case for a very demanding high acuity task), serial search follows and consequently the time needed to locate the target can be very large. A coloured target, on the other hand, is detected in parallel without the need for serial search. This is because adding colour to the target causes a large increase in visual lobe size and this in turn decreases significantly the time required to locate the target. Large objects that in addition to being defined by luminance contrast are also coloured can often be detected and localised in crowded scenes without the need for any eye-movements. In such cases, the visual search is reduced to a single saccade which directs the point of regard onto the target. Large stimuli cannot always be employed since the amount of information to be displayed is substantial in many flight situations as well as in menu bars, usually placed along the sides of the display. More than one saccade is normally required to search for objects in the visual field, but the use of colour reduces significantly the time needed for visual search. Because achromatic attributes are used in the form of text, shape, graphics and shading to represent detailed information, colour appears as a distinct dimension to create conspicuous differences between a target and distractors. Thus, pop-out of colour-coded information in complex scenes is efficient and desirable (Treisman and Gelade 1980). In order to minimise the variability in performance one experiences in natural environments, laboratory experiments often employ visual scenes consisting of distractors that cause visual crowding and a target that differs in one of more visual attributes to the distractor elements. The subject’s task is to search the scene and to locate the target in repeated trials as rapidly as possible. These are the reasons why one of the experiments designed for this study measures how the choice and strength of colour signals employed shortens TCTs and how this advantage is affected in subjects with congenital colour deficiency.

Coding known information by means of colour

While our sensory system can handle a large amount of information, the bottleneck is often our cognitive system when parallel processing of information is not usually the norm.

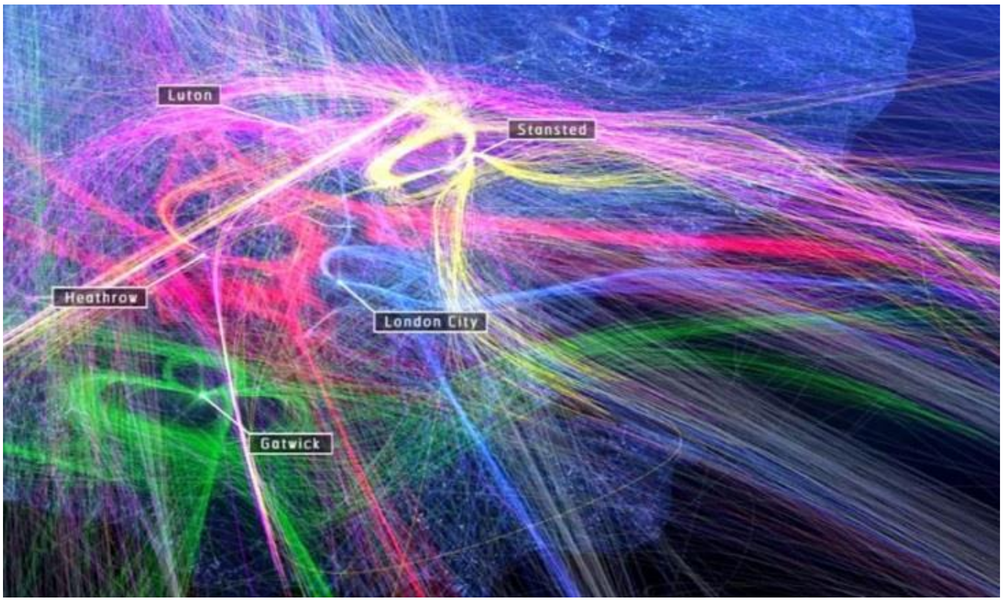

Figure 1: The use of superimposed coloured paths to represent the daily flights to and from each of five UK airports (Paul Haskins, National Air Traffic Services (NATS), 2015).

Source: Daily Mail (hyperlink - http://video.dailymail.co.uk/video/bc/rtmp_uds/1418450360/2015/07/07/1418450360_4342043187001_4341995080001.mp4)

To speed up cognitive processing, one needs to reduce cognitive load by organising complex information into categories defined by visual attributes that can be easily identified and often processed in parallel. One example is the representation of daily flights to and from each of five UK airports as shown in Figure 1. This is made more dramatic by showing the build-up of traffic as a rapidly advancing movie (see link above). A less impressive, but equally important example of the use of colour is the representation of precipitation in ATC displays. Although precipitation varies continuously, it is categorised into six levels in the ATC environment: Levels 1-2 for light weather, Levels 3-4 for moderate heavy weather, and Levels 5-6 for severe weather. Controllers make decisions by identifying the weather levels instead of the precise value of the precipitation. Colour is often used to convey information which is essentially a task of colour naming. Colours that are linked directly to certain objects must be linked correctly to specific colour names. In the example of weather precipitation, the six levels are displayed in different colours. Seeing some areas filled with red, a controller can immediately recognize the presence of severe weather. In the ATC environment, identification of two stimuli is usually performed at separate spatial locations and times. Typically, a controller remembers the colour by its name and searches for and identifies the target by its colour. When certain objects are combined with specific colour signals these objects become easier to remember in identification tasks when memory is required. Moreover, colour becomes increasingly more effective as an aid in the recall of memorized items (Sachtler and Zaidi 1992). For example, when used to identify information such as aircraft shapes, geometric shapes, and alphanumeric signs, colour cues were often more effective than size, brightness, shape or text (Christ 1975). The superiority of colour signals to the use of spatially discrete cues defined by luminance contrast becomes more noticeable in complex scenes when challenging identification tasks are involved.

Segmentation of complex scenes into areas of interest by means of colour

The human visual system organises complex scenes into meaningful objects and / or spatially distinct regions. This is often described as ‘segmentation’ (Pinker 1984). Visual segmentation can enhance performance and make the visual task less demanding and less tiresome. For example, a controller can spatially separate the aircraft situations area, or the number of aircraft of immediate responsibility from the menu areas in a radar display. Thus, when controllers need to find a command in the menu bars, they can direct their attention to specific regions of the display. As ATC displays are usually complex, grouping operations are needed to reduce controller workload. Since the human visual system processes colour separately from achromatic visual features, colour is one of the ways to segment a display into separate regions (Nothdurft 1993).

Segmentation tasks can be either regional or intended to group objects of interest into categories. Regional segmentation involves segmenting a spatially continuous region from its surrounding materials, i.e. filling an area with colour to segment a restricted airspace from non-restricted airspace. Specifically, Yamagishi and Melara (2001) demonstrated that chromaticity information is more effective than luminance in regional segmentation. On the other hand, pattern segmentation involves grouping together spatially discontinuous features that share some common characteristics, i.e. data blocks of aircraft owned by a controller in white and those of un-owned aircraft in green (see Fig. 2). By doing so, the owned aircraft and un-owned data blocks are visually segregated, which in turn makes the visual tasks easier to carry out.

Figure 2: Example of colour coded data blocks used in air traffic control. Data blocks are clusters of small text containing descriptive information of a flight and altitude, the colour indicates whether they are owned or not and whether there are any warnings associated with a particular aircraft.

Source: Heathrow's Advanced Surface Movement Guidance control system.

En-route air traffic control facilities do not require a direct view of the airport or surrounding airspace and the ambient lighting conditions can be controlled. Display operations carried out in the control tower are often more challenging because of varying daylight levels and the possibility of direct sunlight falling on the displays. The luminance and chromatic contrast of the information presented on visual displays can change as a result of variations in ambient illumination (Cardosi and Hannon 1999). The light output of visual displays for use under such conditions should ideally adjust dynamically to account for changes in ambient light level with separate configurations for daytime and night time operation. Glare shields or anti-glare coatings may also be necessary, depending upon the location of the display in the tower.

1.2 Colour vision deficiencies and visual displays

The advancement in the technology of visual displays has provided great opportunities for the use of colour to provide many of the benefits described so far. The obvious requirement is that the operator must be able to make use of colour signals. Subjects with normal trichromatic colour vision possess three distinct classes of cone photoreceptor. These contain short (S), middle (M) and long (L) wavelength sensitive photopigments with appropriate peak absorption wavelengths (λmax). Variant L- and / or M-cone genes can cause significant shifts in the corresponding λmax values and this in turn can cause large changes in chromatic sensitivity. In addition to λmax changes, other factors such as the amount of pigment present in photoreceptors and the relative numbers of L- and M-cones in the retina 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. Colour vision deficiency affects approximately 8% of men and less than 1% of women (Sharpe et al. 1999).

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.						

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

There is little doubt that when the stimulus conditions involve small RG colour differences that are close to normal thresholds, subjects with congenital colour deficiency perform less well when compared with normal trichromats. Studies carried out over several decades have shown that subjects with abnormal colour vision made more errors and produced slower motor responses than those with normal colour vision (Bergman and Duijnhouwer 1980). Mertens and Milburn (1996) examined and compared the performance of normal trichromats and colour deficients on a number of simulations of colour dependent ATC tasks, i.e. colour coding in flight strips, aircraft lights and signal lights in tower operations and colour weather radar. Their results show higher error scores in colour defectives compared to normal trichromats. In addition they found that only two per cent of protanomalous could name the colours (light and dark green, light and dark yellow and light and dark red) of a weather radar display without error, although 43% of deuteranomalous and 10% of deuteranopes could achieve normal scores. Ramaswamy and Hovis (2004) in a study of visual displays used for the control of train movements found that more than half of their 52 colour deficient observers could perform as well as the

99th percentile performance of the colour normal control group in naming eight colours (red, yellow, green, blue green, blue, purple, white and grey) generated on a visual display. Mahon and Jacobs (Mahon and Jacobs 1991) found that observers with abnormal colour vision, except mild deutans who passed the Farnsworth Dichotomous test, D15 (Richmond, USA), made 30 to 40 per cent errors naming the six colours used on aviation electronic flight instrument displays (white, red, green, amber, blue and magenta).

Such findings and significant discrepancies in outcome are not surprising since the severity of colour vision loss in subjects with congenital RG deficiency forms a continuum from almost normal chromatic sensitivity to complete absence of colour vision (Barbur and Rodriguez-Carmona 2012). The level of difficulty associated with colour-related tasks also varies considerably from task to task. In general, redundant cues are almost always involved and the colours used are always well above normal thresholds and stimulate both RG and YB mechanisms. The latter are rarely affected in congenital colour deficiency. Nevertheless, even mild RG colour deficients confuse some colours, particularly those that rely mostly on RG colour differences and this can be of concern in some applications.

1.3 Visual standards for European Class 3 (ATCO) certification

ATCOs are required to identify correctly the colours of aviation lights, and to use effectively the display screen equipment designed for the management of air traffic. Accordingly, ATC applicants must obtain the European Class 3 Air Traffic Controller medical certificate, which requires normal trichromatic colour vision. In spite of this clearly stated requirement, the criteria for obtaining this certificate specifies zero errors on the first 15 plates of the Ishihara test (24 plates edition) which examines only RG colour vision (CAA 2015). In the UK, those failing the Ishihara test (see Section 2.1) will need to take the Colour Assessment and Diagnosis (CAD) test (see Section 2.4) and to pass as ‘normal trichromats’ in order to gain a Class 3 certificate. In other states in Europe, those that fail the Ishihara test can opt to be examined on the anomaloscope (Nagel or equivalent) (EATM 2006), which in turn examines only the applicant’s RG colour vision. The parameters associated with an anomaloscope match show significant inter subject variability and do not correlate well with the severity of colour vision loss (Barbur et al. 2008; Wright 1946). In the UK, the Holmes-Wright lantern type A (HW-A) (see Section 2.3) was accepted as a secondary test when applicants failed the Ishihara plates prior to the introduction of European Class 3 medical requirements.

Follow up colour vision tests are not carried out when the European Class 3 medical certificates are renewed. If the Ishihara plates were used for assessing colour vision, only RG deficiency would be screened for and any yellow-blue loss (either congenital or acquired) would not therefore be picked up. 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 may be appropriate for both RG and YB colour sensitivity to be assessed regularly at medical examinations. These data can then be used to detect when the progression of any inherent (often subclinical) disease yields colour thresholds that fall outside the upper, age-corrected limits established for normal vision (Barbur and Rodriguez-Carmona 2015).

1.4 Problems identified with current and past colour assessment methods

There are two versions of the Ishihara test and the analysis of errors made by normal trichromats and subjects with congenital colour deficiency is specific to the version used. The 38 plates edition includes all the plates that make up the 24 plate version, but the probability of making an error on a particular plate depends on the subject's class of colour vision and is also plate-specific within each class (Rodriguez-Carmona et al. 2012). The analysis presented here is restricted to the use of plates 1 to 15 of the 24 plate version since this version is most commonly used. Before the introduction of the European Class 3 colour vision certification for ATCOs, the standard UK CAA protocol for colour assessment relied on the use of the 24 plates version of the Ishihara test followed by HW-A lantern as a secondary test. All those that failed the Ishihara test were assessed using the HW-A lantern. Table 2 (A, B) describes the expected outcome for this protocol when 1000 applicants are assessed. The calculations are based on the accepted prevalence of congenital colour vision deficiency (see Table 1). Almost all normal trichromats pass the first 15 plates of the 24 plate version with two or less errors (see Table 2A). In addition, 8.3% of deutan subjects also pass when two or less errors are allowed. Those that failed the Ishihara test (see Table 2) were then assessed using the HW-A lantern test. The HW-A lantern is used as a secondary test and all normal trichromats and also 15% of the deutan applicants that failed Ishihara with two or less errors also pass. If all the deutans are tested on the HW-A lantern, 22% pass. These results confirm earlier findings which have been known for some time and stated by the CIE (CIE 2001). What is important to note is that the protocol based on first 15 plates of the Ishihara 24-plate version with two or less errors as a pass followed by the HW-A lantern test passed all normal trichromats and 22% of applicants with deutan deficiency. About 8% of applicants were required to carry out the secondary test.

Predicted outcome per 1000 applicants (Ishihara pass: 2 or less errors)				
Applicants	**1000**	**No. that fail Ishihara**	**No. that fail HW-A**	**No. of ATCs**
Normals	920	4	0	920
Deutans	60	55	47	13
Protans	20	20	20	0
Total	**1000**	**79**	**67**	**933**
% of applicants that undergo secondary tests=				**7.9**
% of normal that fail Ishihara =				0.5
% of deutans that pass =				21.7
% of protans that pass =				0.0
% total colour deficient subjects that pass =				16.3

Table 2a: Predicted outcome per thousand applicants based on two or less errors on plates 1 to 15 of the Ishihara 24 plates test as a pass followed by the HW-A lantern test.

Predicted outcome per 1000 applicants (Ishihara pass: zero errors)				
Applicants	**1000**	**No. that fail Ishihara**	**No. that fail HW-A**	**No. of ATCs**
Normals	920	92	0	920
Deutans	60	59	47	13
Protans	20	20	20	0
Total	**1000**	**171**	**67**	**933**
% of applicants that undergo secondary tests =				**17.1**
% of normal that fail Ishihara =				10.0
% of deutans that pass =				21.7
% of protans that pass =				0.0
% total colour deficient subjects that pass =				16.3

Table 2b: Predicted outcome per thousand applicants based on zero errors on plates 1 to 15 of the Ishihara 24 plates test as a pass followed by the HW-A lantern test. The percentages of applicants that fail either Ishihara or HW-A lantern within each class are based on results obtained in 742 subjects (Rodriguez-Carmona et al, 2012). The least affected deutans who pass Ishihara with either zero errors or two or less errors would also pass the HW-A lantern test. Since the overall outcome is the same, the only advantage of using Ishihara as the primary test is to reduce significantly the number of applicants that require a lantern test.

The first protocol was followed for decades and ensured that all normal trichromats and ~ 22% of deutan subjects passed and could therefore obtain the medical certification required to work as ATCOs. EU Commission Regulation (No 1178/2011) requires ATCO applicants to have normal trichromatic colour vision which is most commonly assessed using a criterion of zero errors on the first 15 plates of the Ishihara 24 plates test version.

When the protocol is changed so that a pass requires zero errors, the number of colour deficient subjects that pass the full protocol (with the HW-A lantern as a secondary test) remains the same, but 10% of normal trichromats and almost all subjects with congenital colour deficiency fail the Ishihara test and go on to do the secondary test. The latter becomes essential if one wishes to ensure that all subjects with normal trichromatic colour vision pass. If the HW-A lantern test is not carried out, the EU regulation requiring that all ATCO applicants that pass should have normal trichromatic colour vision is achieved (at least in respect of RG colour vision by allowing zero errors on the Ishihara test). More importantly, the unwanted consequence is that 10% of normal trichromats will not pass. The HW-A lantern is no longer being manufactured and existing lanterns have been unserviceable for several years. To overcome this problem, UK CAA have replaced the HW-A lantern with the CAD test. Figure 3 shows the small but clear separation between the least sensitive normal trichromats and mildest deuteranomalous subjects. The CAD test also examined YB colour vision.

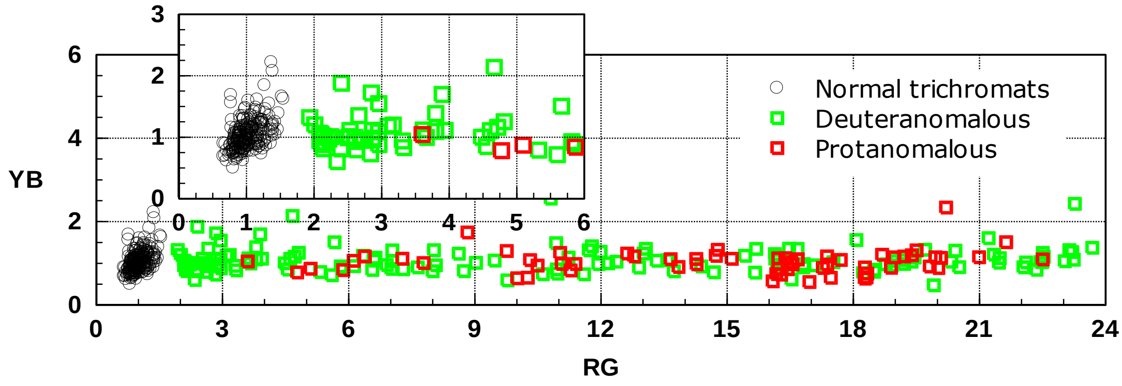

Figure 3: RG and YB CAD thresholds measured in normal trichromats (n=333) and deutan (n=269) and protan (n=132) subjects. The results show the clear separation between the least affected deuteranomalous subjects and the least sensitive normal trichromats.

Since 2009 the CAA have been using zero errors on plates 1-15 of the 24-plate version as the pass criterion. The severity of colour vision loss for those that fail the Ishihara test is then quantified using the CAD test (Civil Aviation Authority 2013). ATC applicants are required to have normal trichromatic colour vision. The CAD test has the sensitivity and specificity required to diagnose correctly the class of colour vision involved and to quantify the severity of YB and RG loss. This protocol ensures that all ATC applicants that have normal trichromatic colour vision pass. The same protocol is also used to assess pilot applicants, but different pass / fail limits are employed depending on the applicant's class of colour vision deficiency (UK CAA and FAA 2009). As a result of applying the new CAD based pass / fail limits to pilot applicants, ~ 35% of subjects with congenital colour deficiency pass and are classed as safe to fly.

When using the full CAD test, the subject's RG and YB thresholds are measured and compared against the age-matched upper normal limits. The subjects are also classified automatically as:

- Normal trichromats,
- Deutan-like,
- Protan-like,
- Tritan-like,
- Acquired loss of colour vision.

The strict adherence to current regulations for ATC operators (i.e., normal trichromatic colour vision) means that the 22% of deutans that passed secondary testing prior to 2009 and were allowed to work as ATCOs in the past can no longer do so. Although these deutan-like observers may not perform colour-related, ‘threshold’ tasks as well as subjects with normal trichromatic colour vision, they are nevertheless able to cope with the ‘suprathreshold’ use of colour in the ATC environment and can name accurately the white, red and green signal colours in the size and varying intensities employed in the HW-A lantern.

This clear example justifies the need to make best use of past evidence and current knowledge to ensure that ATCO applicants with mild colour deficiency are not disadvantaged unfairly by requiring normal trichromatic colour vision. It is generally accepted that the aim of regulatory authorities is to ensure that colour deficient applicants who pass ‘specified requirements’ are able to achieve the level of performance in safety-critical ATC tasks that may reasonably be expected of subjects with normal trichromatic colour vision. There are at least three possible options that should be considered:

- The need to conduct a detailed study to assess the use of colour signals within the ATC environment, similar to what was carried out for professional pilots (UK CAA and FAA 2009). Set up pass / fail limits based on the evidence that is likely to emerge from this study. This option will be discussed further in section 3.
- Restrict the chromaticities of the colours employed in ATC applications to those that generate large YB colour signal differences and set a RG pass / fail limit similar to that employed for pilots (i.e., thresholds < 6 CAD units). These measures would ensure that adequate, residual colour discrimination remains and that the supra-threshold colours employed will always be detected by those that pass.
- Develop and adopt a practical approach using modern colour assessment tests with pass / fail limits that virtually replicate the outcome of earlier practices based on Ishihara and the HW-A lantern. This approach would ensure that those that pass can name accurately very small red, green and white signal lights of varying intensity pass. The HW-A limit based on

past practices does not rule out a higher limit being equally appropriate, but in the absence of a detailed study, no solid evidence exists at this time to justify a higher limit. This option will be discussed further in section 4 of this report.

There are also further considerations that justify the need to establish safe, minimum colour vision requirements that are appropriate within specific environments and to avoid the easier alternative (from a regulatory viewpoint) of requiring every applicant to have normal trichromatic colour vision. The recent UK Disability Discrimination Act (2004) has to a certain extent exposed weaknesses in the current standards and procedures. Companies rely on existing regulations and require applicants to hold a valid medical certificate that conforms to these regulations. The problem arises when these regulations are not adequate and reliable evidence can be produced to demonstrate that a colour deficient applicant is able to carry out essential, colour-related occupational tasks with the accuracy and efficiency that can be expected of normal trichromats.

Chapter 2

Description of the relevant colour vision tests

2.1 Ishihara plate test

The Ishihara pseudoisochromatic plates test consists of a series of numbers outlined by different coloured dots as shown in Figure 4. This is the most widely accepted screening test for RG colour deficiency and uses camouflage to exploit the expected colour confusions of colour deficient observers (Belcher et al. 1958; Birch 1997; Frey 1958; Sloan and Habel 1956). 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 the visual task, plates 2-15 for screening, plates 16-17 for protan / deutan classification and plates 18-24 contain pathways which are intended for the examination of non-verbal subjects (only the first 15 plates are used for screening in aviation). The Ishihara test employs a range of designs, such as transformation, vanishing or hidden digit. In the vanishing type plate (Fig. 4B) a figure is seen by colour normals but not by colour 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. 4A), with careful placement of the colour dots giving an apparent transformation of the perceived figure; normal trichromats see one figure and colour deficient people see a different figure in the same design. Positive evidence of colour deficiency is given by transformation designs whereas vanishing designs give negative evidence. The test employs luminance and YB colour noise and is therefore limited to RG deficiency. The test does not assess loss of YB sensitivity.

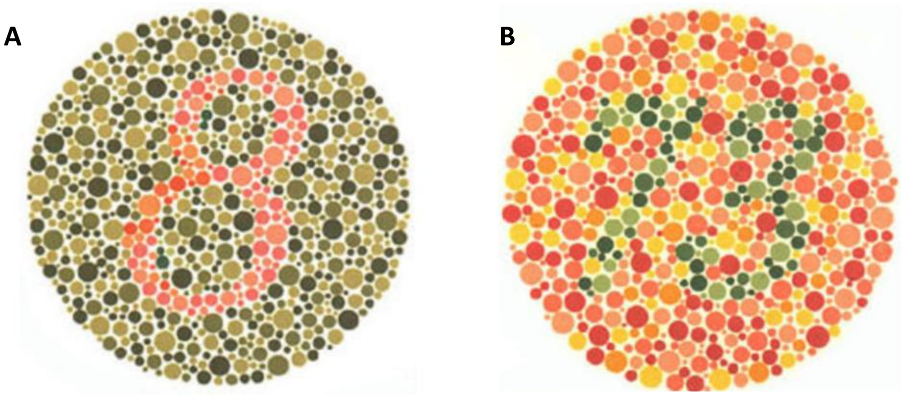

Figure 4: Examples of the types of plates used in the Ishihara test (24 plate edition): (A) a transformation plate; while a normal trichromat would see the number '8', colour deficients would see '3' (B) a vanishing plate; the number '73' would not be visible to most colour deficients.

The plates are viewed at about 65-75cm (i.e., an arm's length) distance using a MacBeth easel lamp for illumination. The book is placed in the tray beneath the lamp which has been filtered to have a spectral power distribution that is equivalent to average daylight. The direction of illumination is approximately 45^{o} with respect to the surface of the plate. The choice of illuminant used is important because the reflectance of the patches on the plates have been chosen to have certain chromaticities when illuminated with daylight. The examiner instructs the person being tested to report the number they see on each plate 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 colour deficient subjects should see this number. With a viewing time of no more than 3 seconds allowed for each plate, undue hesitation on the part of the subject is the first indication of colour deficiency.

Figure 5 shows that 90.2% (213) of the subjects with normal trichromatic colour vision make no errors on the first 15 plates of the 24-plate version of the Ishihara test and almost all normals (except for 1) get all plates correct with 3 or less errors. Figure 5 shows that the probability of making k or less errors is much greater for deutan than for protan subjects. If the number of Ishihara plates failed is a valid indicator of the severity of colour vision loss then the results of Figure 5 demonstrate that for the same number of errors made, the severity of colour vision loss is much greater in protan than deutan subjects. For example, 26% of deutan subjects make 8 or less errors compared with only 7% of protan subjects. Note that no subject made errors on the introductory plate, hence the maximum number of possible errors is 14.

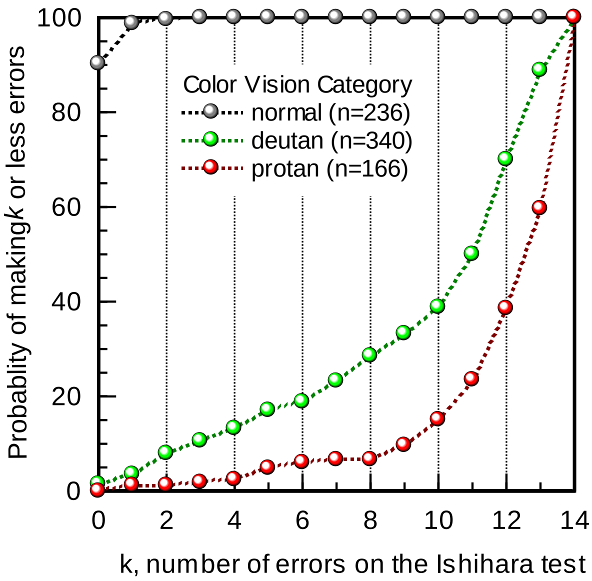

Figure 5: The probability of making *k* or less errors when reading the numerals on the first 15 plates of the Ishihara 24-plate test version plotted for a group of normal trichromats and for subjects with congenital, deutan, and protan colour deficiency.

2.2 Nagel anomaloscope

The Nagel anomaloscope (Fig. 6) is based on colour matching and is often regarded as the standard clinical reference test for identifying and diagnosing red / green colour deficiency, 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 colour mixture ratio in the top field, and the other to alter the luminance of the yellow lower field (as shown on the right in Fig. 6). The test is administered in two stages. Usually only the dominant eye is fully tested and the other eye is then checked to ensure the same match. This confirms the high probability of any congenital colour deficiency. Following familiarisation 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 colour and brightness. The subject is not asked to name the colours. 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 minimise the effect of chromatic after images, the subject looks away from the instrument into the dimly lit room for a few seconds and then 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 has to just alter the luminance of the lower yellow half of the field and see if an exact “match” in both colour 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 colour vision; whether colour deficiency is deutan or protan; and whether the subject is a dichromat

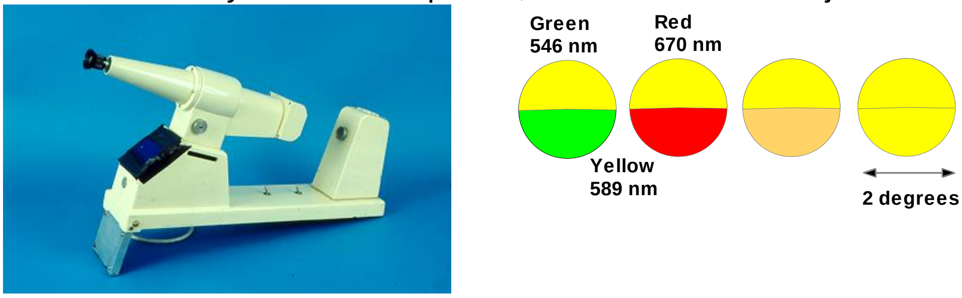

Figure 6: Photograph of the Nagel anomaloscope (Model I, Schmidt and Haensch, Germany) and schematic illustration of the Nagel anomaloscope split field. The percentage mixture of red to green in the top half and the luminance of the yellow bottom field can be changed until a match of the two fields is achieved.

(absence of a cone-type) or an anomalous trichromat (anomalous cone-type). The matching range and the match midpoint should also provide an indication of the severity of the RG colour vision loss. Normal Nagel match parameters have however been recorded in a small number of subjects with clear loss of RG sensitivity as assessed using the CAD test. Such observations and the large variation in match midpoint in ‘normal’ trichromats have not passed unnoticed in earlier studies. “*While the anomaloscope is thus eminently satisfactory for distinguishing the protanomalous from the deuteranomalous observer, the correlation between the abnormality of the yellow match and the deficiency of colour discrimination is not always maintained. Although on average, the observer with poorer colour discrimination makes a more erroneous yellow match than the observer with better discrimination, the reverse can happen with individual observers*” (quotation from W.D.Wright 1946). More recently a model based on the genetic analysis of cone pigment genes in congenital colour deficiency was developed to examine how spectral shifts in the wavelength of peak spectral sensitivity and the corresponding optical densities of L- and M-cones, as well as their relative numbers in the retina affect the parameters of the yellow match (Barbur et al. 2008). The model predicts the shifts in midpoint caused by differences in optical density in subjects with normal trichromatic vision and also the normal match parameters observed in some subjects with significant loss of RG chromatic sensitivity. The latter observation is predicted by the model in subjects with variant L- and M-cone pigments who make relatively normal matches, but exhibit reduced RG chromatic sensitivity (Barbur et al. 2008). These observations and the need for an experienced examiner make the Nagel anomaloscope less attractive, particularly when some level of colour deficiency can be allowed as a pass.

2.3 Holmes-Wright lantern (type A) (HW-A)

The first colour vision lanterns were devised in the late 19th Century and were introduced as vocational tests for occupations as a practical means of determining whether applicants can identify signals and navigational aids.

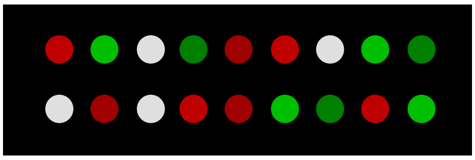

Figure 7: The nine pairs of vertical lights that are presented to an applicant as part of the HW-A colour screening test.

The Holmes-Wright lantern type A (see Fig. 8), manufactured in 1974 and was designed to reproduce the spectral features of some earlier lanterns, but with improved mechanical construction and modern light sources (Holmes and Wright 1982). HW-A lantern is currently one of the CIE recommended colour vision tests for the transport services (CIE 2001). Although still is use, the HW-A lantern is no longer manufactured or supported.

The HW-A lantern shows two vertical colours (Fig. 7 & 8A). Two reds, two greens and one white are used, which have x, y chromaticity co-ordinates within the internationally agreed specifications for signal lights (CIE 2001). Nine pairs of the colours are shown representing all the possible colour combinations (Fig. 7). The lanterns are viewed at 6 m (20 ft). Before beginning the test, the examiner demonstrates (the 'DEM' setting shown in Fig. 8B) the colours by showing red, green and white lights in the top positions of the first three pairs and naming them correctly. The examination is then carried out, starting in fairly dim room illumination, with colours at high luminous intensity (200 µcd) ('HIGH' setting in Fig. 8B).

In order to assess variability, several repeats of the HW-A test were carried out as part of this study. The nine pairs of colours were shown three times, making 27 presentations of about five seconds each. The subject had to name the top and then the lower colour shown each time. The room was then darkened and the subject dark adapted for about 15 minutes; the test was then repeated again with the three sequences of nine colour pairs. Thus each subject was shown 54 pairs of lights. The examiner noted each colour misnaming the subject made. The results were sorted into number of types of misnaming for both light levels, e.g. 3 greens called 'white', 2 whites called 'green', etc. The protocol followed by the CAA (prior to the introduction of the CAD test) consisted firstly in showing one run and if all nine pairs of lights were named correctly, no further runs were carried out and the subject passed the test (CP3 or 'colour-safe'). If there were any errors (excluding red / green confusions), two more runs were carried out. In order to pass the test (CP3), the subject had to report correctly the colours of each of the 18 pairs of lights involved in the last two runs. However, if the subject made one or more errors, one more test was carried out in the dark following 15 minutes of dark adaptation. If the subject made no errors on this final run, the results were taken as a pass (CP3); otherwise the subject was considered 'colour-unsafe' or CP4. If at any stage the subject named a red as 'green' or a green as 'red', then the test was discontinued with a CP4 certification outcome.

A B

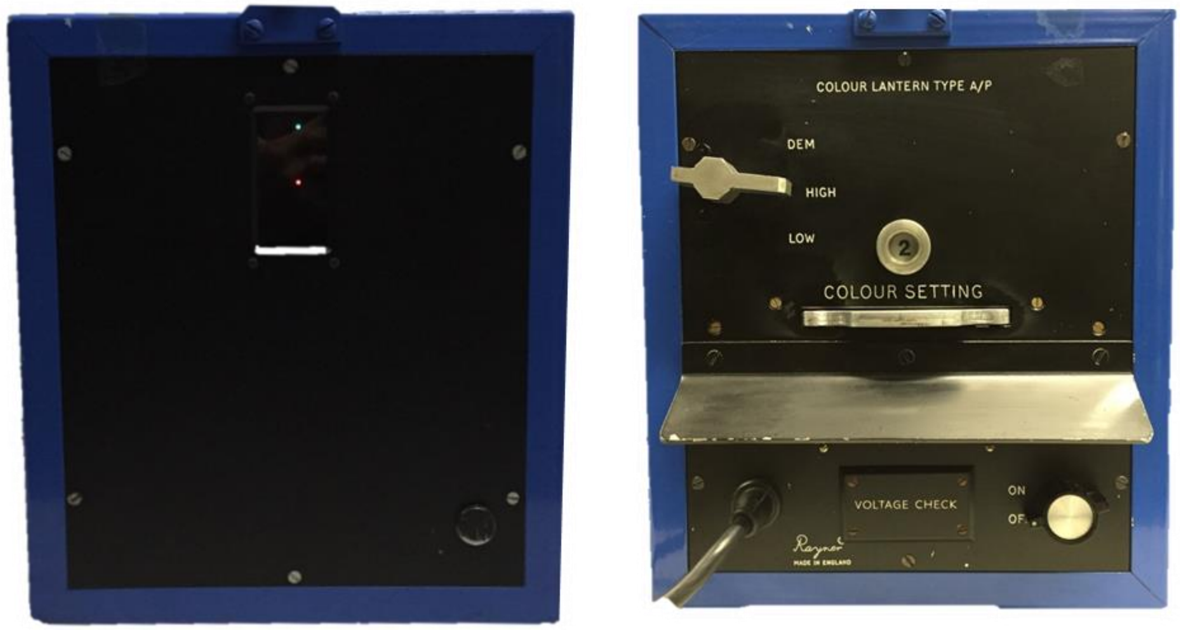

Figure 8: Photographs of the Holmes-Wright type A lantern; front view (A) and rear view (B).

2.4 The CAD test

The Colour Assessment Diagnosis (CAD) test has been described in an earlier CAA report (CAA 2009; CAA 2006). The CAD test is implemented on a calibrated visual display and consists of coloured 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 (see Fig. 9B). The subject’s task is to report the direction of motion of the colour-defined stimulus by pressing one of four appropriate buttons. Randomly interleaved staircase procedures are used to adjust the strength of the colour signals involved according to the subject’s responses in order to determine the thresholds for colour detection in each direction of interest. This makes it possible to establish reliable estimates of red-green and yellow-blue colour thresholds. The CAD test has a number of advantages over conventional tests both in terms of isolation of colour signals as well as sensitivity and accuracy:

- Isolation of colour signals is achieved by masking luminance contrast signals generated by the moving coloured stimulus that is only photopically ‘isoluminant’ for the standard CIE ‘normal’ observer. This is particularly important since there is a large variation in L:M cone ratio within normal trichromats (Carroll et al. 2002) and the variation in cone spectral responsivity functions in colour deficient observers will introduce variations in the perceived luminance contrast of most coloured stimuli. This is simply because the resulting luminance efficiency function, V(λ), is likely to vary both amongst normal trichromats and within colour deficient observers. When using dynamic luminance contrast masking the applicants cannot make use of any other cues apart from colour to see the moving target and to carry out the task.

- The severity of both red-green (RG) and yellow-blue (YB) colour vision loss is quantified in Standard Normal Units (SNU) which are easy to understand (see Fig. 10A&B) (Barbur et al. 2006; Rodriguez-Carmona et al. 2005).
- The CAD test has close to 100% sensitivity and specificity in detecting congenital colour deficiencies and in classifying the type of deficiency involved. There is also excellent agreement (Barbur and Rodriguez-Carmona 2012) with the Nagel anomaloscope (with a kappa statistic of 0.97 based on 289 subjects). Since subjects with two potential variant genes can produce normal anomaloscope matched, but exhibit RG colour thresholds just outside normal limits (Barbur et al. 2008), there may well be good reasons why 100% agreement with the anomaloscope cannot be achieved. In terms of sensitivity, the CAD test can detect even minimal colour deficiencies (particularly in subjects with acquired loss of chromatic sensitivity) that may produce variable results or pass unnoticed in conventional colour vision tests.
- The availability of built in, monocular and binocular normal, upper threshold limits from 6 to 85 years of age (see Fig. 11A&B) make the CAD test particularly useful in allowing for normal aging changes and in diagnosing acquired loss of RG and YB colour vision.

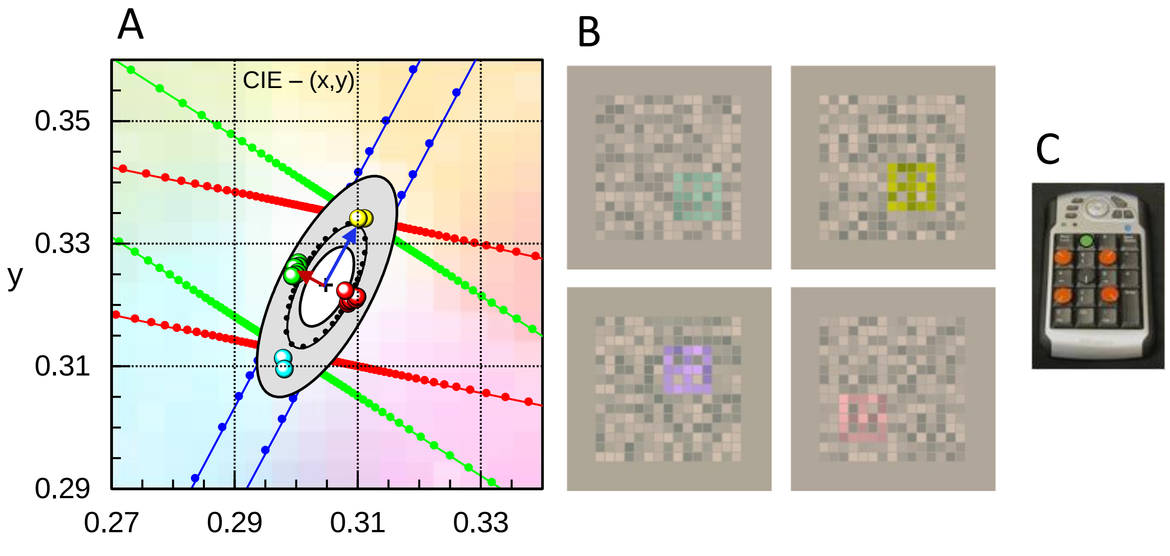

Figure 9: (A) The statistical limits for the standard normal (SN) CAD observer are plotted in the CIE (x,y) 1931 chromaticity chart. The black cross at the centre of the diagram shows the chromaticity of the white background, (xb, yb): 0.305, 0.323. The dotted black ellipse represents the mean values computed from the distribution of RG and YB thresholds in 333 normal trichromats. The red and blue arrows indicate the corresponding SN unit for RG and YB colour vision, respectively. The grey-shaded area represents the statistical distribution of thresholds in a young population of normal trichromats. The inner and outer ellipses represent the 2.5% and the 97.5% limits of variability, respectively. The red, green, and blue dotted lines denote the "colour confusion bands" based on data measured in protanopes, deuteranopes, and tritanopes, respectively. The large coloured dots within the centre grey area plot measured thresholds that are typical of a normal trichromat. (B) Screen dumps showing the RG and YB stimuli employed in the CAD test. (C) Bespoke numeric keypad with raised buttons used by the subject to indicate the direction of movement of the coloured stimulus.

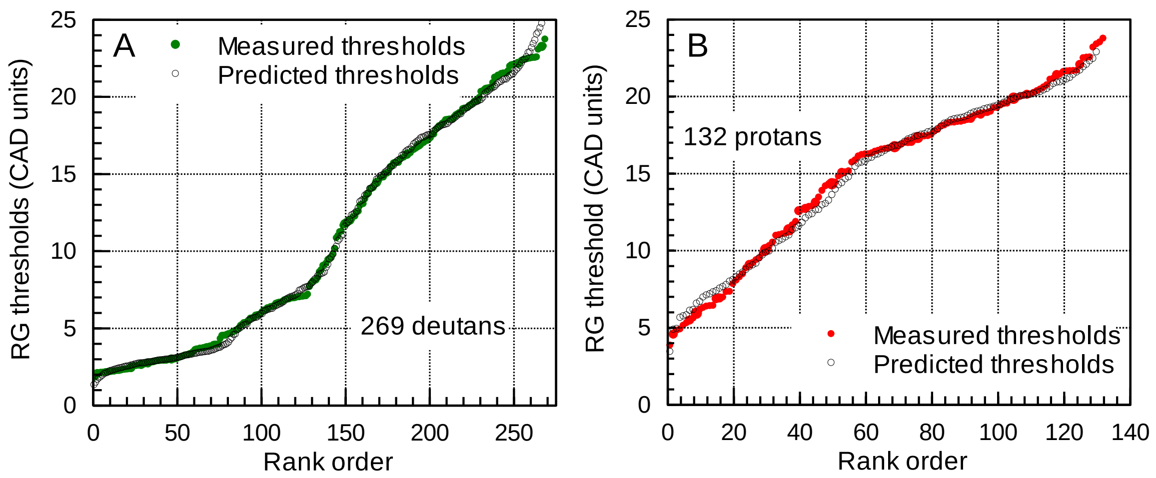

Figure 10: Ranked distribution of RG colour thresholds measured in subjects with congenital (A) deutan- and (B) protan-like deficiency. The spread in RG thresholds varies over a large range; mildest deutan deficiency starts at a RG threshold of just over 2 SNU whilst the lowest protan deficient subject measured had a threshold of ~4 SNU. The black circles show the best prediction of the ranks based on samples taken from four discrete distributions in the case of deutan-like subjects (i.e., four distinct subgroups that are needed to describe deutan deficiency) and only two to three discrete distributions in the case of protan-like subjects (Barbur and Rodriguez-Carmona 2012).

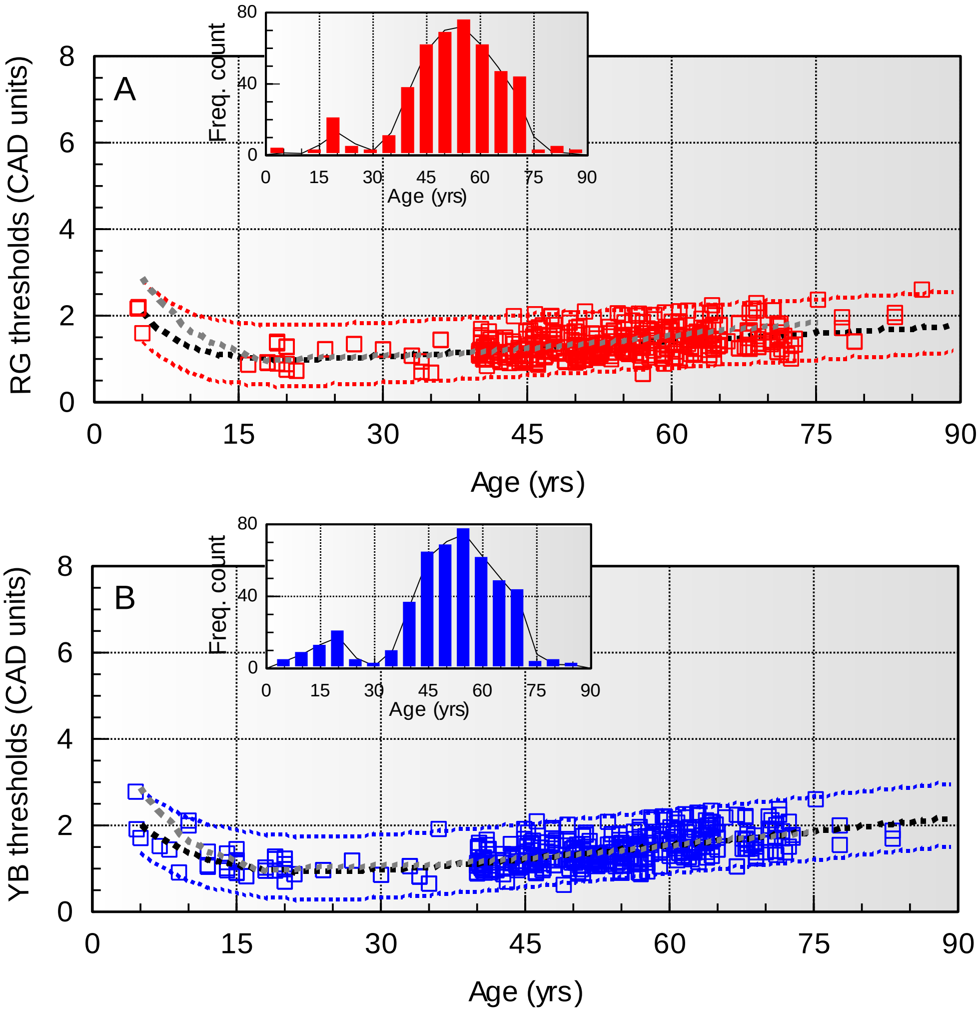

Figure 11: The statistical limits of RG (A) and YB (B) normal variability as a function of age for a sample of normal eyes. The dotted lines represent the mean thresholds and the corresponding ±2.5σ limits as a function of age (Barbur and Rodriguez-Carmona 2015). The results reveal the well-documented increase in thresholds below 10 years of age (Knoblauch et al. 2001) the optimum age for best thresholds around 20 years, and the remarkably gradual and linear increase which almost doubles both the RG and the YB thresholds during the normal life span. The age dependence of CAD thresholds is also in excellent agreement with the square root of the total error score (√ TES) as measured in 382 normal subjects on the F–M 100-hue test (Kinnear and Sahraie 2002). The latter are shown by the dotted grey line which is almost indistinguishable from the mean CAD thresholds above 15 years of age.

Chapter 3

Setting minimum colour vision requirements

3.1 Identification of the most safety-critical colour vision tasks in ATC

The exact benefits of colour vision in occupational environments are often difficult to assess, largely because of redundancy in the coding of visual information. The task of setting minimum colour vision requirements that are both fair and safe remains an important challenge, but the studies carried out for pilots and train drivers and the establishment of pass / fail limits based on task-related levels of performance equivalent to those measured in normal trichromats may offer an acceptable alternative to less rigorous practices that are currently employed in many occupational environments.

The methodology introduced by the CAA (UK CAA and FAA 2009) to establish minimum colour vision requirements for pilots based on the applicant's ability to carry out the most demanding, colour related tasks with the same accuracy as normal trichromats, can also be applied to other professional environments. This method can be summarised as follows:

- The first task is to establish the most important, visually-demanding, colour-related tasks when the use of colour can enhance significantly the performance the operator can achieve over longer times, or / and the correct processing of colour signal involves safety-critical tasks.
- Produce accurate (photometrically, radiometrically and spatially equivalent) simulations of the most demanding tasks in the laboratory and develop a method for quantifying the subject's performance in these tests.
- Measure percentage correct scores on each task identified as important or / and safety-critical in normal trichromats and also in subjects with varying severity of colour vision loss.

- Assess the severity of colour vision loss in each congenital colour deficient subject included in the study using a test that quantifies accurately the loss of chromatic sensitivity.
- Relate the subjects' ability to perform the specified tasks (in terms of efficiency and accuracy) to their colour thresholds and use these findings to establish pass / fail limits which ensure safe and acceptable levels of performance.

Although many colour-related tasks can be identified within the ATC environment, the most important involve the use of large-field visual displays and the correct interpretation of signal lights (as seen from control towers). Many colour-related tasks do not require a direct view of the airport or surrounding airspace (particularly in an en-route control environment), and so ambient lighting conditions can be controlled.

Work in control towers may also require the detection and correct interpretation of signal lights and the use of colour displays in ambient environments with poorer control of ambient illumination. Under such conditions, adequate display designs that make optimum use of visual parameters such as background luminance, object size, colour and luminance contrast become more important (Cardosi and Hannon 1999). In addition, there are no internationally or nationally agreed standards or specification regarding display and related human factors aspects. The use of colour in ATC displays is only one of the many aspects related to how information is displayed. Facilities often exist to change the colour palette within the air traffic management software to suit individual Air Navigation Service Provider requirements.

The majority of colour-related tasks in the ATC environment involve large-field visual search where targets of interest presented on a display and located amongst multiple distractors must be processed efficiently. Based on the fact that any colours could potentially be selected to convey warnings related to a particular flight, the most critical condition in terms of colour deficient use would be when coloured targets separated by small chromatic differences fall along deutan or protan confusion axes.

In view of such observations, the work in this study focused on experiments designed to quantify the advantages of colour vision in ATC displays, the enhanced performance that can be achieved in optimum designs and the extent to which improved designs can benefit subjects with anomalous trichromatic colour vision. In addition to examining and relating the severity of colour vision loss to accuracy and speed of performance in large-field visual displays, we also examined how subjects with the same range of colour vision loss perform tasks that involve identification of signal lights. The measures of efficiency and accuracy achieved on display tasks and the percentage correct scores achieved on lantern tests were then compared against measures of the subject's RG and YB chromatic sensitivity. This approach made it possible to grade the severity of colour vision loss into functional categories that may be appropriate for use within different occupational environments.

3.2 Visual search

The effect of reduced RG chromatic sensitivity on visual search times when colour signals are used in addition to other cues has been investigated previously (Cole et al. 2004; Cole and Macdonald 1988; O'Brien et al. 2002). The findings from these studies show that colour deficients require significantly longer visual search times to complete the task. The selection of parameters for the visual search task in these studies was not however intended to reflect the demands on the use of colour in ATC displays and the performance achieved in normal trichromats and colour deficient subjects was not related directly to accurate estimates of their RG and YB chromatic sensitivity.

Performance in visual search is generally quantified by measuring the time taken to locate an object (i.e., the target of interest) that differs in some visual attribute to a large number of similar objects, usually described as distractors. The target can differ from distractors in one or more visual attributes such as size, orientation, contrast, colour or it can be flickering or moving. The time taken to locate the target stimulus amongst distractors relates strongly to the 'visual conspicuity' of the object (Barbur et al. 2003; Walkey et al. 2005) and is taken to reflect the capacity of the visual system to extend the visual field over which the target feature of interest is processed in parallel. In addition to measuring visual search times, it is also important to check the percentage of correct scores achieved by the subject and this usually requires a second response to select one of a possible number of alternatives. Our visual search experiments required two responses from the subject. The first response signalled the detection of the target. When this button is pressed, the visual display returns to a uniform field. The program then waits until the subject indicates the location of a gap in a ring or its colour using a four-alternative, forced-response procedure. A variant of this experiment employed 36 achromatic distractors (all Landolt rings with randomly selected gap orientation), three coloured distractors and a coloured test target. The subject's task is to search for and locate the Landolt ring shown in one of four specified colours. Once the target is located, the subject has to press the centre button on the numeric keypad (see Fig. 13) to indicate the detection of the target. The pressing of the button returns the screen to the uniform background and the program waits for the subject's second response which is needed to indicate the orientation of the gap or the colour of the target. In this way, it is possible to measure the subject's response accuracy, by simply measuring the percentage of correct responses during the test.

3.3 CRATO test study

The Colour Requirements for Air Traffic Operators (CRATO) test was designed to measure the mean time needed to carry out the task, labelled as 'Task Completion Time' (TCT) and the 'Correct Response Scores' (CRS) for visual tasks that involve large visual fields and for stimuli of equivalent size and contrast to the data blocks employed in ATC displays. Although the objects employed are somewhat abstract since they consist of similar targets and distractors, the use of such stimuli makes it possible to evaluate how target contrast and colour affect visual performance. Stimuli were presented over a square region subtending ~ 20° of visual angle on a high resolution, 'spectraview', NEC monitor (PA301W, Tokyo, Japan). The background field was set at a luminance of 32 cd/m2 and had a chromaticity of xb = 0.305, yb = 0.323 in CIE – (x,y) 1931 chromaticity chart. The coloured stimuli in all experiments were defined as chromatic displacements from background chromaticity (xb, yb) in specific colour directions. The programs needed for the study were developed by City Occupational Ltd (London, UK). The display calibration programs were the same as those employed in the CAD system (Barbur and Connolly 2011).

As part of this study 70 subjects were examined: 33 normal trichromats and 37 subjects with deutan- and protan-like colour deficiencies. The age of the subjects ranged from 17 to 65 years (mean 37 years, median 34 years). All subjects had best corrected visual acuities of 6/9 or better.

Stimuli developed for CRATO experiments

The following stimulus designs have been produced in order to quantify how luminance and colour contrast affect TCTs and CRSs in visual search tasks. We started by establishing the relationship between target luminance contrast, correct scores and TCTs on achromatic displays when spatial features are used to distinguish the target from distractors. The stimuli developed to achieve this aim consist of achromatic 'ring' distractors and a single target ring with a gap that can take one of four randomly assigned locations (Fig. 12).

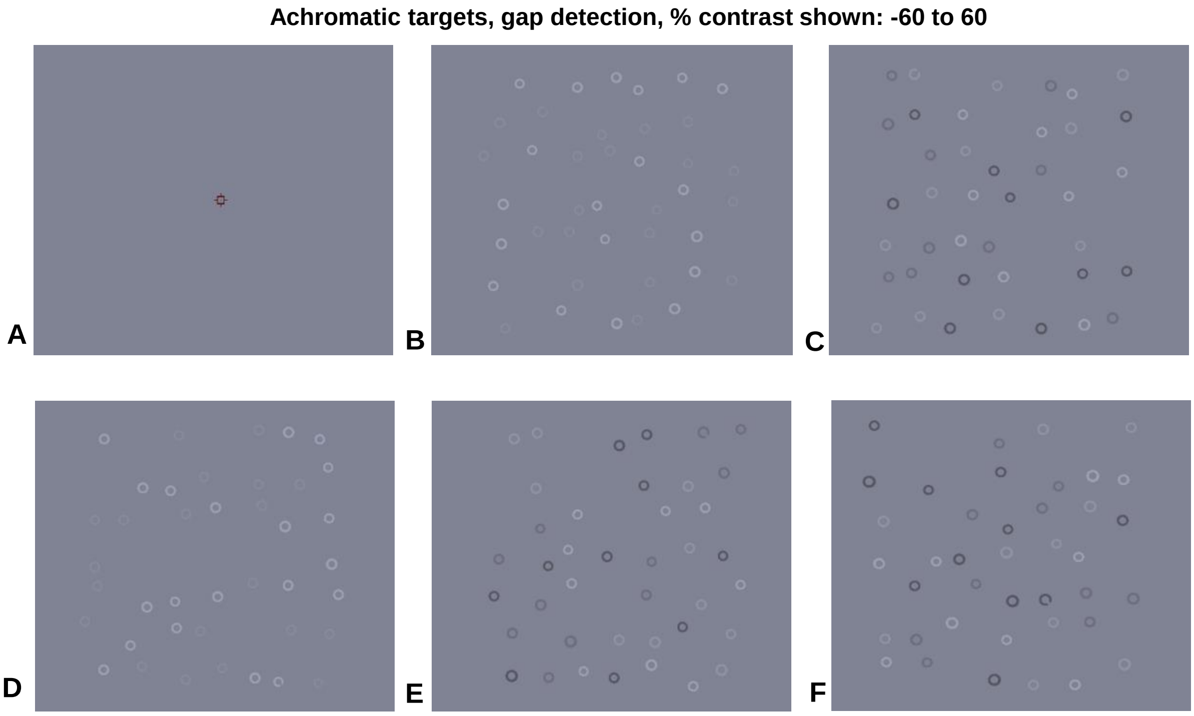

Figure 12: Examples of displays designed to investigate the effect of luminance contrast on visual search. Section (A) shows the uniform field with the centre cross that guided the participant's point of regard before stimulus onset. The remaining sections show visual search scenes with targets of 15% (B), 30% (C), 60% (D), -30% (E) and -60% (F) contrast.

The gap size equals the width of the ring which in turn equals 1/5 of the outer ring diameter. The rings can have either negative or positive contrast with gap sizes selected randomly in the range 5 to 6 min arc. When viewed directly, the subjects have no difficulty in resolving the gap since its size is approximately equivalent to a full letter size at the limit of resolution of the eye (i.e., ~ 1'). 40 distractors were usually employed.

The outer diameter of targets and distractors were selected randomly within the range 25' to 30'. Data blocks on ATC displays are not of a fixed size and some variation will occur based on the amount of information available for a particular flight (see Fig. 1). The choice of size accounts for the fact that data blocks are made up of multiple characters and the visual acuity of ATCOs is presumed to be normal (6/6 when tested binocularly).

Prior to each presentation, a central fixation is displayed briefly to attract the subject's point of regard to the centre of the field. Each target specification (i.e., each colour saturation or luminance contrast) was presented to the subject ~90 times so as to be able to compute the mean TCT and CRS performance parameters. If a subject failed to respond within 12 seconds from the start of a presentation, then that presentation was terminated and recorded as an error.

Class I experiment – Achromatic visual search

A fixation target (as shown in Fig. 14, section A) is presented briefly for ~ 350 ms in the centre of the screen. This is done to guide the subject's point of regard to the centre of the field. Some 500 ms after the fixation is turned off, the display is populated with randomly distributed distractors. The target is positioned randomly and consists of a ring with a gap. Up to four different luminance contrasts can be investigated in the same experiment by randomly allocating one of four contrasts to the test target. The subject's task is to search for and locate the target. One of four possible gap locations (i.e., top right, top left, bottom right and bottom left) is selected randomly in each display. The subject's first task is to press the centre button on the numeric keypad as soon as the target is located. This action returns the screen to a uniform background. The subject's second task is to press one of the four 'red' buttons (see Fig. 13) on the numeric keypad to indicate the location of the gap. This is an intuitively obvious task that requires negligible learning to carry out. On average, the time needed to locate the target and to note the orientation of the gap (i.e., the task completion time) decreases with increasing target contrast. Although some differences exist between positive and negative luminance contrast, both reach an asymptote above ~ 60% contrast (Fig. 13).

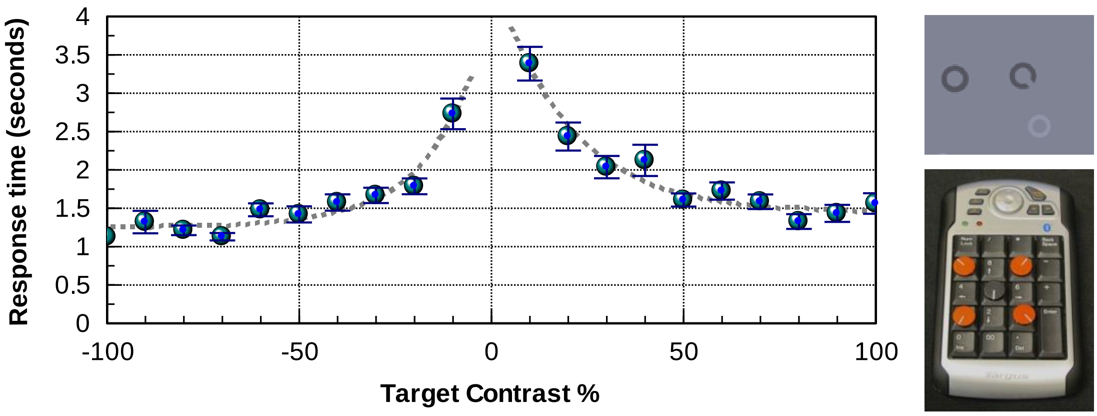

Figure 13: Typical results showing the effect of target luminance contrast on task completion times (TCT). The subject's first task was to locate the target and to register the location of the gap. When the luminance of the target approaches that of the background, i.e. the target approaches 0% luminance contrast, the task can no longer be carried out, in fact below 10% contrast measurements took too long to complete. Little or no improvement in TCTs can be observed beyond ~ ±60% contrast. CRS values in this test were close to 100%.

Class II experiment – Single coloured target with achromatic distractors

In these experiments the distractors are defined only by luminance contrast and can be either full rings or Landolt Cs with randomly selected gap positions. The target stimulus was defined by both luminance and colour contrast. When full ring distractors were employed with a Landolt C target, the task could be carried out in the absence of colour signals (Fig 14 A,B&C). This we label as **Class IIA** experiment. The stimulus colour became very important when Landolt Cs were used as distractors (see Fig14 D,E&F). In the absence of detectable colour signals, the task could not be carried out. This condition we label as **Class IIB** experiment. The two stimulus conditions distinguish between the redundant (full ring distractors and Landolt C target) and the non-redundant (all Landolt Cs) use of colour signals. TCTs were measured for a number of chromatic displacement directions that generate colour contrast signals in only RG, only YB or RG and YB chromatic mechanisms. A number of colour signal strengths (i.e., chromatic saturations) were investigated along each direction. Figure 14 shows typical single coloured targets employed in Class II TCT experiments. The addition of colour signals causes what is often described as 'pop-out' and this shortens considerably the TCTs that can be achieved with only spatial cues.

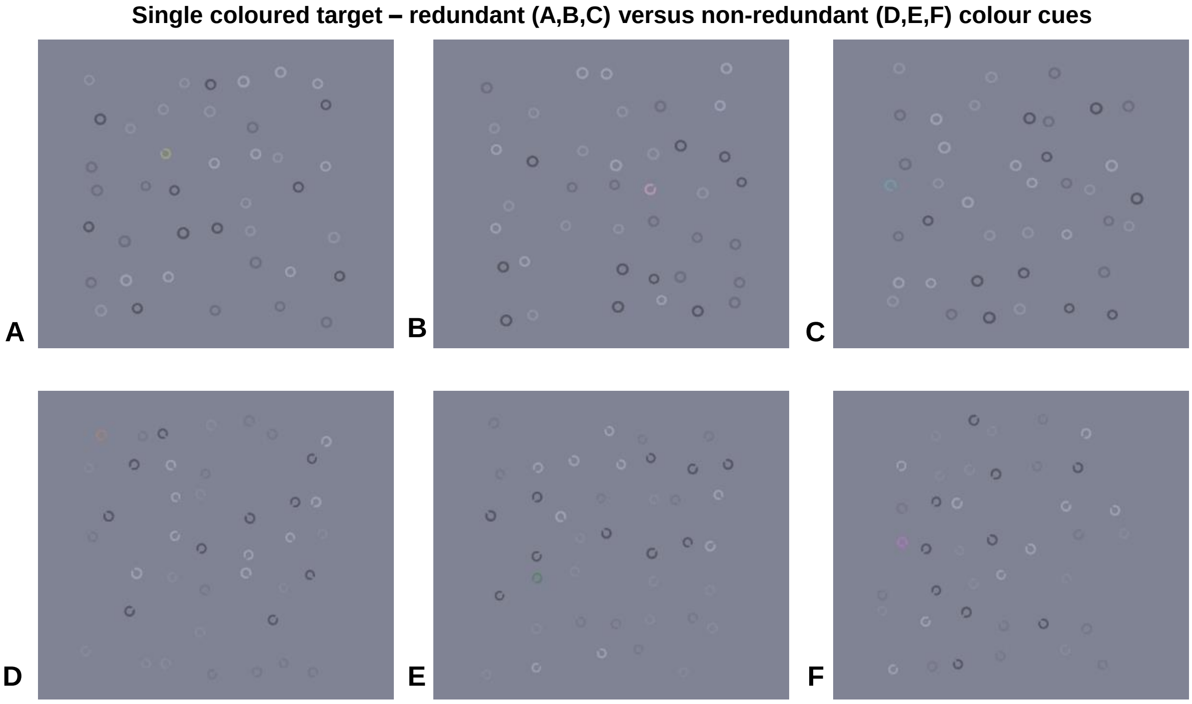

Figure 14: Examples of displays designed to investigate how colour can reduce TCTs in visual search tasks. The target in sections A, B and C is defined by colour as well as the presence of the gap. Colour is therefore used redundantly. The target shown in sections D, E and F can only be distinguished from the background distractors by means of its colour signal. When the chromatic saturation is large, the target 'pops-out' amongst distractors and the spatial cue becomes redundant. When this happens the visual performance is determined entirely by the colour cue. As a result, suprathreshold chromatic saturations, yield TCTs that are statistically equivalent for the two conditions (i.e., the presence of the gap adds no additional advantage to TCTs).

Another variant of this (i.e., a **Class IIC**) experiment also used only colour to define the target (i.e., both target and distractors were shown as full rings so as to eliminate the spatial cue, see inset to Fig. 15B). In order to measure the percentage correct response scores, the subject was required to name the colour of the target. Four test colours were interleaved randomly in each experiment and the single coloured target displayed could have one of four possible colours which were shown to the subject before the start of the experiment. The subject's first task was to press the target detection button as soon as the target was located. The display screen returned within a few milliseconds to a uniform field and the subject's remaining task was to name the colour of the target by pressing one of four buttons assigned to yellow, blue, green and red colours. Since colour was the only target cue, TCTs became much longer as the chromatic saturation approached the subject's chromatic threshold for the colour direction investigated. Surprisingly, this was less so for 'blue' targets which were easier to detect in the periphery of the visual field. This variant task is in principle similar to the Class IIB experiment since the target is defined only by colour. The task was however more demanding since in addition to discriminating between four possible colours, the subject also had to remember the colour assignments of the four buttons (See inset to Fig. 15B). Typical data in experiments using Class IIA and Class IIC are shown in Figure 15 for a 60yrs old subject with normal trichromatic colour vision. Although the results are very similar when the colours employed are well above threshold (which is usually the case in ATC displays), the absence of the spatial cue in the Class IIC experiment results in longer TCTs when faint colours are employed. The majority of experiments that involved single coloured targets were therefore carried using either Class IIA (colour and spatial cues) or Class IIB (colour only cues) experiments.

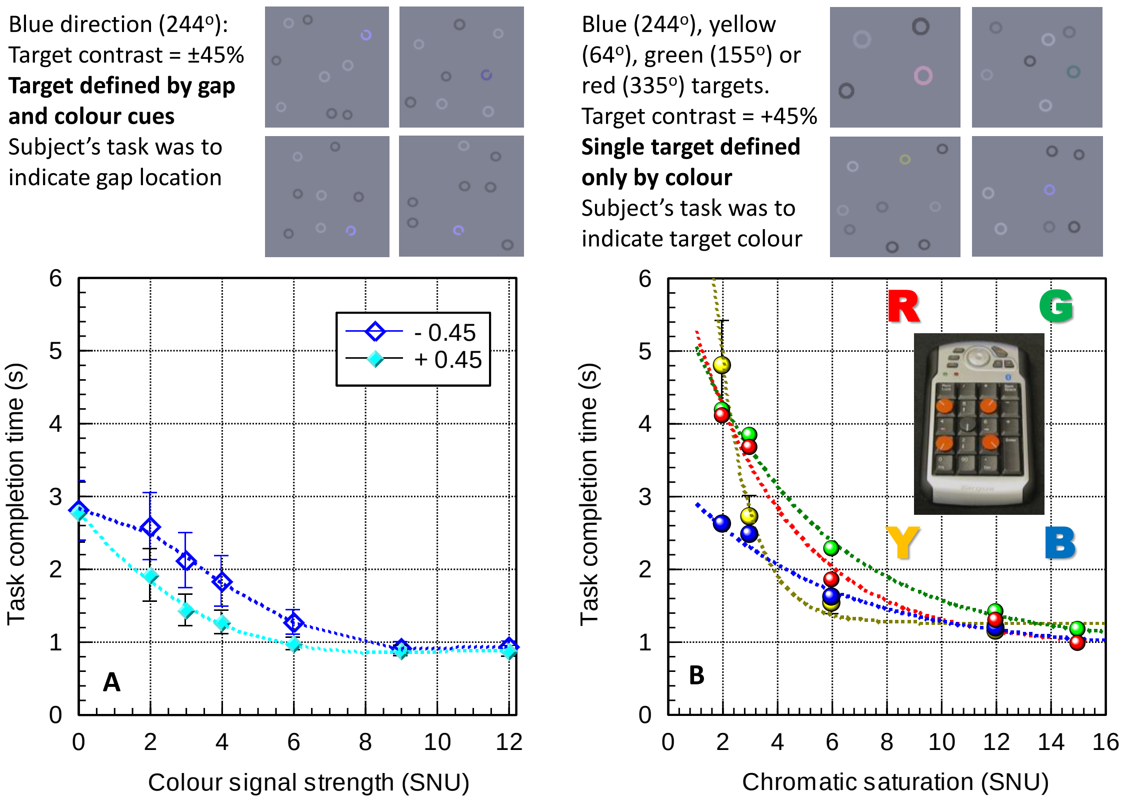

Figure 15: Example of TCTs measured using Class IIA experiments (i.e., target defined by gap and colour cues, section A) and full rings (in the absence of spatial cues, section B) when the task of the subject was to name the colour of the target (Class IIC). Although colour signals are more effective for targets of positive luminance contrast, the TCTs converge to similar values for both positive and negative contrast when the colour signal is absent (A). Blue and yellow colours appear to be marginally more effective in large visual fields at low chromatic saturations, but the benefit of colour signals becomes very similar for all colours when the chromatic saturation is large.

Class III experiment – Multiple coloured objects with achromatic distractors

In addition to achromatic distractors and the coloured target, three additional distractors were also coloured. Four different coloured objects were therefore always present in each display, as shown in Figure 16. As in Class I and II experiments, one of four luminance contrasts were allocated randomly to each object. This experiment is important since the use of more than one colour in a complex display may increase the need for trichromatic colour vision in order to benefit fully from the use of colour signals. The use of colour and spatial gap (Class IIIA) or colour only (Class IIIB) cue to define the test target gave rise to two distinct experiments. In Class IIIA experiments, only rings were used to generate both coloured and achromatic distractors. The subject had to search the scene for the Landolt C target which was always presented in the same colour, but could have one of four, randomly selected luminance contrasts. The target could therefore be identified by its colour and / or the presence of the gap. The test procedure remained the same in all experiments. The subject’s first task was to press the centre button on the numeric keypad as soon as the target was located. This caused the display screen to return to a uniform field and the subject’s second task was to press one of four corner buttons to indicate the orientation of the gap. Class IIIB experiments were again very similar, but the spatial cue was made ineffective by adding a randomly selected gap location to each ring in the display. The only cue available to locate the target was therefore its colour. The subject was shown all four colours at the beginning of the experiment and was then instructed to search for one of the four colours. Percentage correct responses were again measured by requiring the subject to press one of four buttons to indicate the location of the gap in the test target after each presentation. Although for very small chromatic signal strengths (that were close to the subject’s colour detection threshold), the two types of experiment yield vastly different TCTs, when larger colour saturations were employed, the two methods yielded very similar TCTs which suggests that colour coding is the dominant cue, even when the display contains multiple colours. As a result of these observations, the majority of experiments were carried out using Class III type experiments.

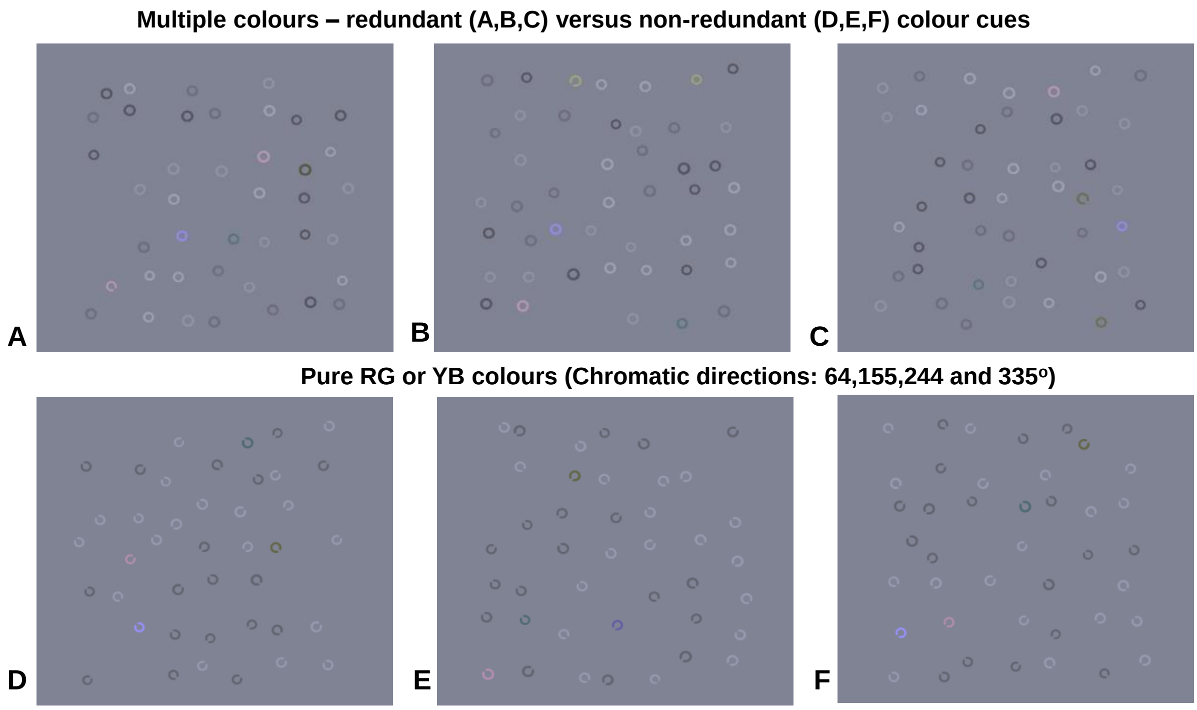

Figure 16: Examples of Class III displays designed to investigate how the use of multiple coloured objects affects TCTs in visual search tasks. The target in sections A, B and C is defined by colour as well as the presence of the gap. Colour is therefore used redundantly. In sections D, E and F the target has no distinguishable spatial cues and consequently its detection relies entirely on the correct processing of its colour. In practice, suprathreshold chromatic saturations, yield TCTs that are statistically equivalent (i.e., the presence of the saturated colour is sufficient to achieve the shortest possible TCT).

Class II experiments (colour 'pop-out') – Preliminary findings

To reveal the advantages of the use of colour in visual search an initial study was carried out to investigate by how much observers benefit from 'pop-out', i.e. when colo ur allows for rapid separation of target objects from distractors, see Figure 17.

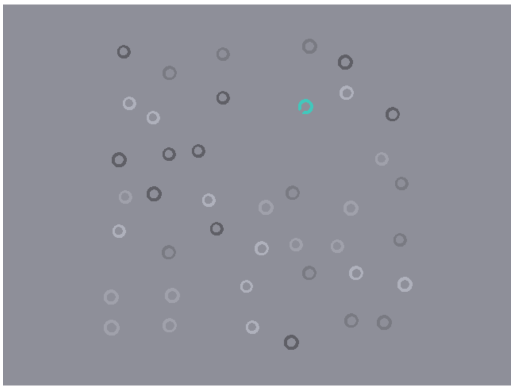

Figure 17: Reminder of CRATO pop-out experiment (Class IIA) with a coloured target defined by a large chromatic displacement of 12 SN CAD units (i.e., 12 times above the mean normal threshold) towards the 'green' region of the spectrum locus. In the absence of colour, a TCT of ~ 3s is needed on average to locate the target. With pop-out the, TCT is well under 1s. Following target detection and response, the subject's second task is to report the location of the gap in the ring (in this case by pressing the bottom left button on the numeric keypad).

Figure 13 shows the response times when the target was not defined by colour, but instead by a spatial cue; the target was a Landolt C whilst the distractors were complete rings. These results provide a baseline measurement which allows us to evaluate the advantage of pop-out. When colour is added to the target visual search times are improved. Figure 18 shows that the addition of yellow (62°), blue (242°), green (157°) and red (337°) separately, lowered the response times compared with the achromatic target (which corresponds to a chromatic displacement of zero CAD units). Targets and distractors had a percentage luminance contrast of ±60% and ±30% with respect to the uniform background. These values are typical of luminance contrasts employed in ATC displays. TCTs were measured for chromatic saturations from 2 up to 28 CAD units in the RG direction and 2 to 12 CAD units, along the YB axis. This range of chromatic saturations is limited by the maximum limits that can be achieved on the visual display along the colour confusion lines, away from the selected background chromaticity. The presence of the additional spatial cue (i.e., a Landolt C target versus full ring distractors) meant that the subject can carry out the task, even in the absence of a colour signal. The addition of a colour signal facilitates the visual search process with TCTs below one second for larger chromatic saturations. In addition, yellow and blue targets appear to illicit faster responses for equivalent saturations. The polarity of luminance contrast for coloured targets does not appear to have a significant effect on TCTs. For all four coloured targets, there was no effect of chromatic displacement or luminance contrast with virtually 100% correct scores.

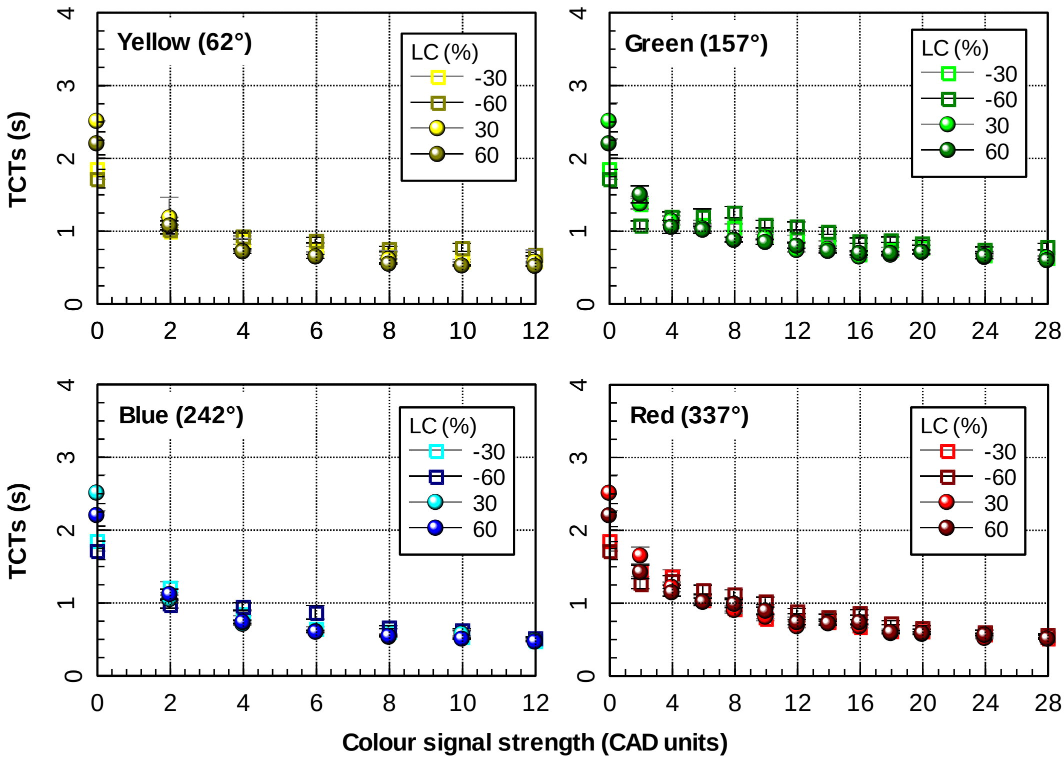

Figure 18: Class IIA experiment – pop-out. TCTs measured in a young subject with normal trichromatic colour vision as a function of colour signal strength measured along each of four directions away from the background chromaticity in the CIE 1931 chromaticity chart. The task is to locate a coloured target (Landolt C) amongst 40 achromatic distractors (rings). Achromatic response times to the same target are shown at a chromatic displacement of 0 CAD units. Data points represent the average response times of 90 presentations with error bars representing the standard error.

Effective luminance contrast of coloured targets in colour deficients

TCTs measured as a function of colour signal strength in protanomalous subjects produced unexpected findings which merit further discussion. The luminance contrast of the coloured targets (as seen by a subject with normal trichromatic colour vision) in all CRATO experiments was always kept constant at a fixed value and independent of chromatic signal strength. In this way, we were able to investigate the benefit of increasing chromatic saturation on task completion times independently on luminance contrast. We know from the results obtained in Class I experiments that luminance contrast can affect TCTs, particularly for low contrasts. The results in subjects with normal trichromatic colour vision show that TCTs decrease to an asymptote with increasing colour signal strength or in the case of Class I experiments with increasing target luminance contrast. In the case of normal trichromats, the presence of the additional spatial cue makes it possible to identify the target and to report the location of the gap correctly for every contrast value and chromatic saturation investigated.

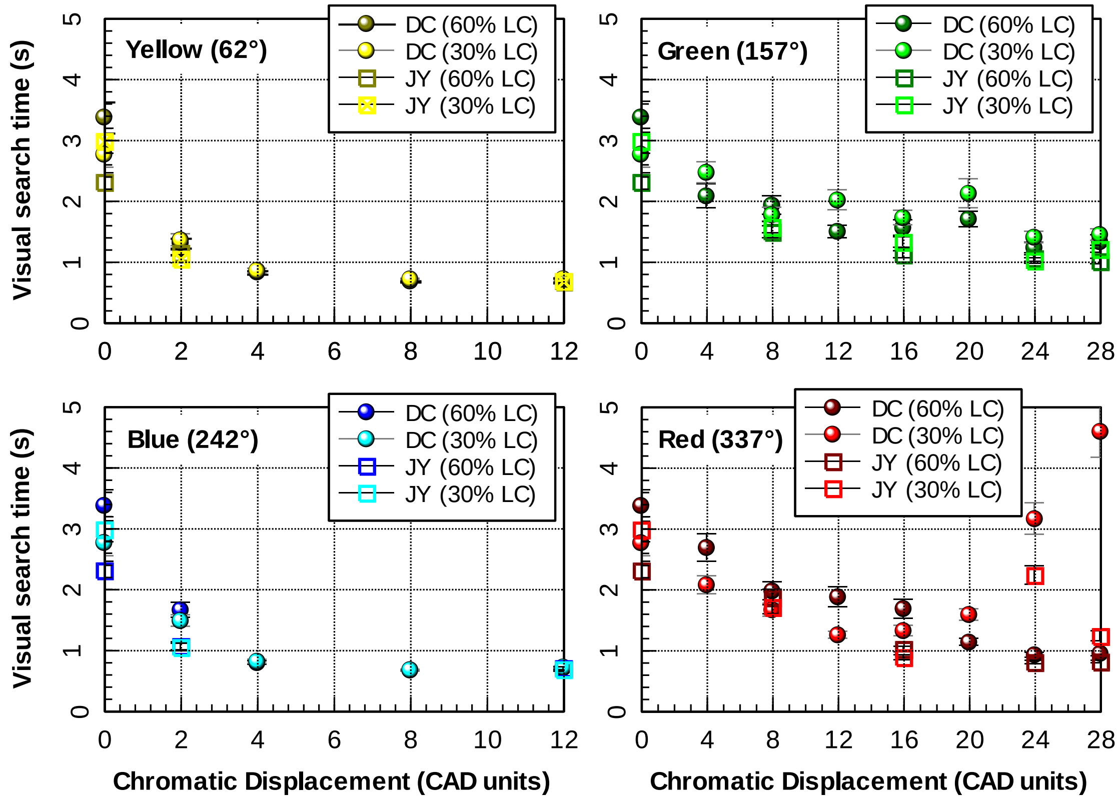

Figure 19a: Class IIA experiment. TCTs for two protan observers for yellow, green, blue and red targets as a function of colour signal strength (in CAD units). The targets had luminance contrasts of +60% and +30% and could therefore be detected easily by the colour deficient subjects when no colour signal was present. Although the luminance contrast of the targets remains independent of colour signal strength in normal trichromats, this is not the case in subjects with congenital deficiency and this can affect the 'conspicuity' of the target and hence the TCTs. Subject DC and JY had RG CAD thresholds of 6.8 and 21.5 SN units, respectively. The average CRS value measured in these tests was 99.2 ± 1.8%.

TCTs for protanomalous subjects (Fig.19A and B) improve with chromatic saturation for yellow, blue and green target colours, but the opposite can be the case when one starts with targets of positive contrast and the increase in colour signal is towards the red region of the spectrum locus. This effect can be accounted for by the predicted changes in effective luminance contrast and the weak RG chromatic signal expected in a subject with protan deficiency. The right combination of initial target luminance and chromatic signal strength can reduce significantly the effective luminance contrast of the target in subjects with protanomaly or protanopia. This prediction also applies to targets of negative contrast and chromatic displacements towards the 'green' region of the spectrum locus. The precise outcome remains difficult to predict accurately since the effectiveness of the residual RG colour signal may also depend on the luminance contrast of the target. One can however say with confidence that in the absence of a strong chromatic and / or luminance contrast signal, the target can no longer be detected efficiently and this inevitably leads to long TCTs, as observed experimentally. In the absence of a strong chromatic signal, the subject relies on the luminance contrast to locate the ring with the gap, but the latter is being either increased or decreased by increasing the chromatic saturation of the target. It is therefore not unexpected to find that for some colour deficient subjects, the initial luminance contrast of the target can be reduced significantly by the addition of colour contrast. The subject can still rely on the use of some colour signals and in principle an initial positive contrast can also become negative by increasing the colour signal strength. These possibilities may account for the increased TCTs and the unexpected changes measured for large chromatic saturations in these subjects. The effect is also diminished or absent when the direction of chromatic displacement does not correspond to the protanopic colour confusion line and the increased chromatic saturation also generates YB colour difference signals which the subject can use to carry out the task.

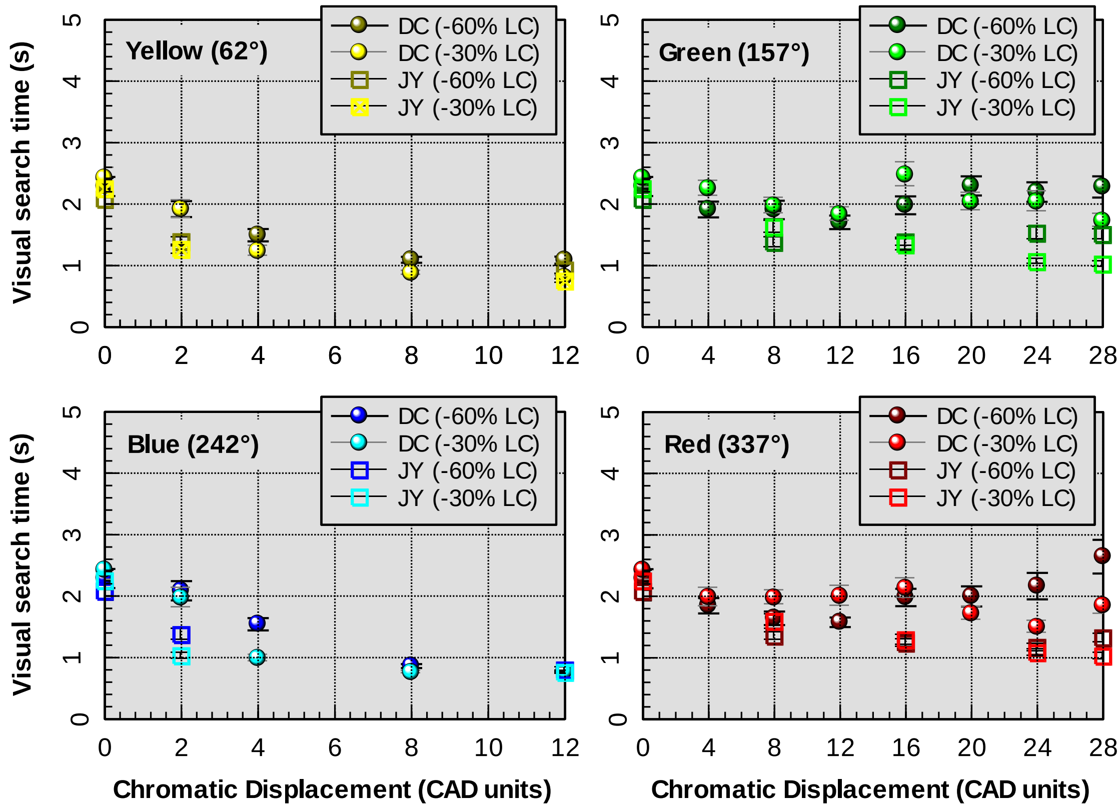

Figure 19b: Class IIA experiment. TCTs measured as a function of colour signal strength in two protan observers for yellow, green, blue and red targets defined by luminance contrasts of -60% and -30%. Subject DC and JY had CAD thresholds of 6.8 and 21.5 SN units, respectively. The average CRS value measured in these tests was 99.4 ± 1.0%.

Deutan subjects show improvement in TCTs with increasing target saturation for both red and green target colours (Fig. 20A and B). This is more obvious for the mildly affected deutan subject (JL), whereas for the more severe subject (LS) this improvement is less obvious, with only a slight reduction in TCTs for targets of largest chromatic saturation.

For yellow and blue targets, both protan and deutan subjects, performed the visual search task with the speed and accuracy measured in normal trichromats.

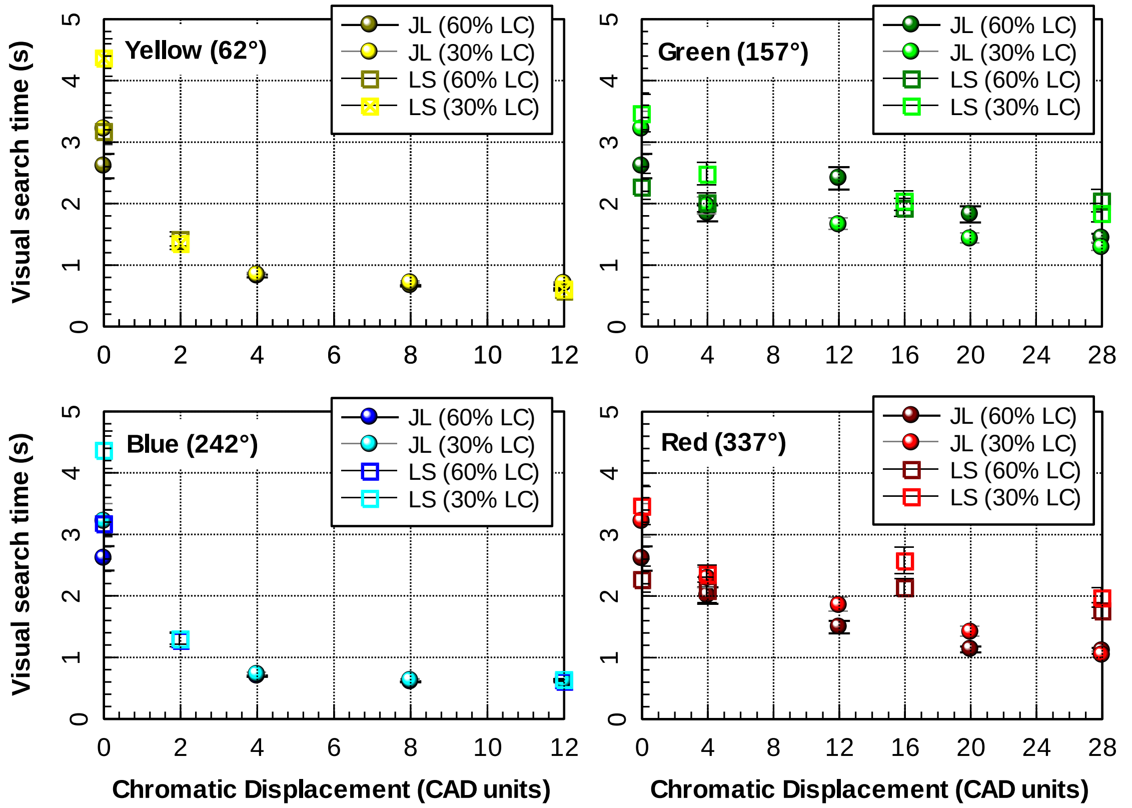

Figure 20a: Class IIA experiment – pop-out. TCTs measured in two deutan observers for yellow, green, blue and red targets with positive luminance contrasts of 60% and 30%. Subject JL is a mild deutan with a CAD threshold of 3.4 SN units, and LS is a severe deutan with a threshold of 19.5 SN units. The average CRS value measured in these tests was 99.5 ± 0.9%.

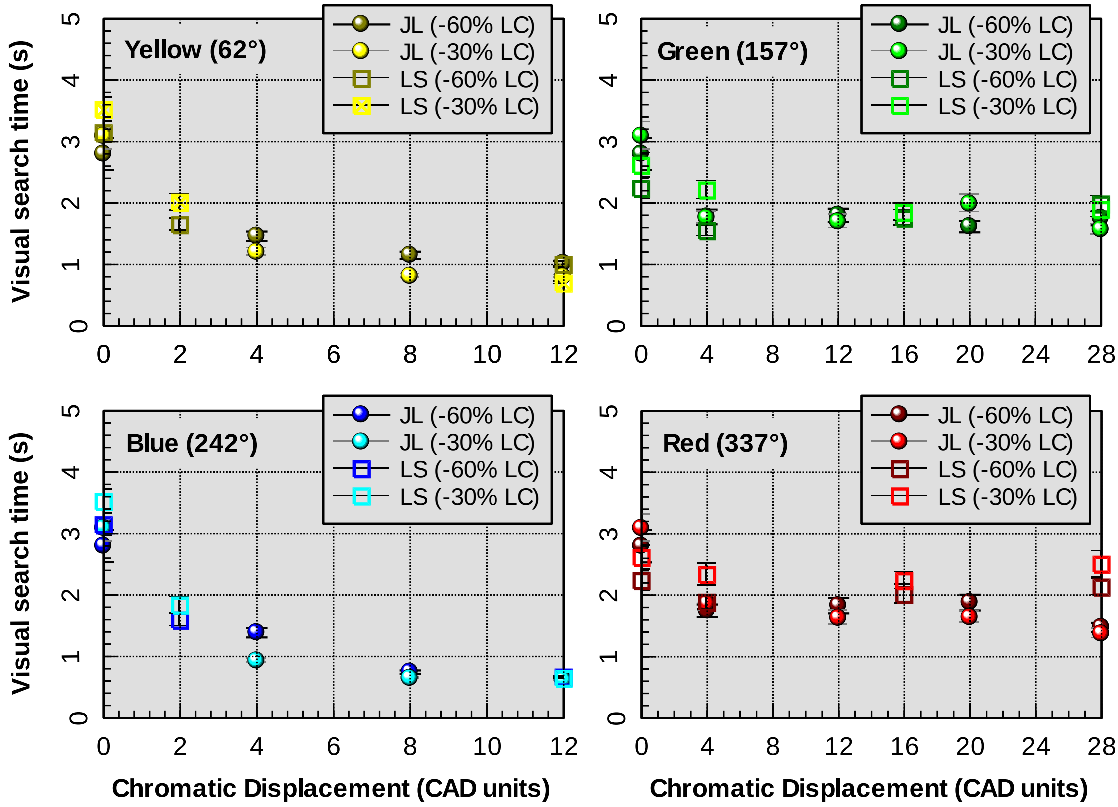

Figure 20b: Class IIA experiment – pop-out. TCTs measured in two deutan observers for yellow, green, blue and red targets with negative luminance contrasts of 60% and 30% compared to the background field. Subject JL is a mild deutan with a CAD threshold of 3.4 SN units, and LS is a severe deutan with a threshold of 19.5 SN units. The average CRS value measured in these tests was 99.6 ± 0.8%.

Effect of eccentricity across a visual display

Large field displays are often used in ATC applications to present the complex information needed to carry out the task. Coloured objects are often seen in the periphery of the visual field, but the colours appear less saturated, particularly when red or green. Yellow and blue targets, on the other hand, are somewhat less affected. As a result, both normal trichromats and subjects with congenital RG colour deficiency perform well in large-field, visual search tasks when yellow or blue colour signals are added to a target defined by luminance contrast. The effect is particularly strong for targets of positive contrast. A separate experiment was carried out to investigate the extent to which the better performance achieved with yellow and blue targets may, at least in part, be linked to the differential loss of RG and YB chromatic sensitivity with eccentricity. The results in Figure 21 show how YB and RG thresholds vary with target eccentricity for stimuli defined by luminance contrast. Yellow-blue detection thresholds show only a small increase whereas red-green thresholds increase more rapidly with a significantly greater loss of chromatic sensitivity at 10° eccentricity.

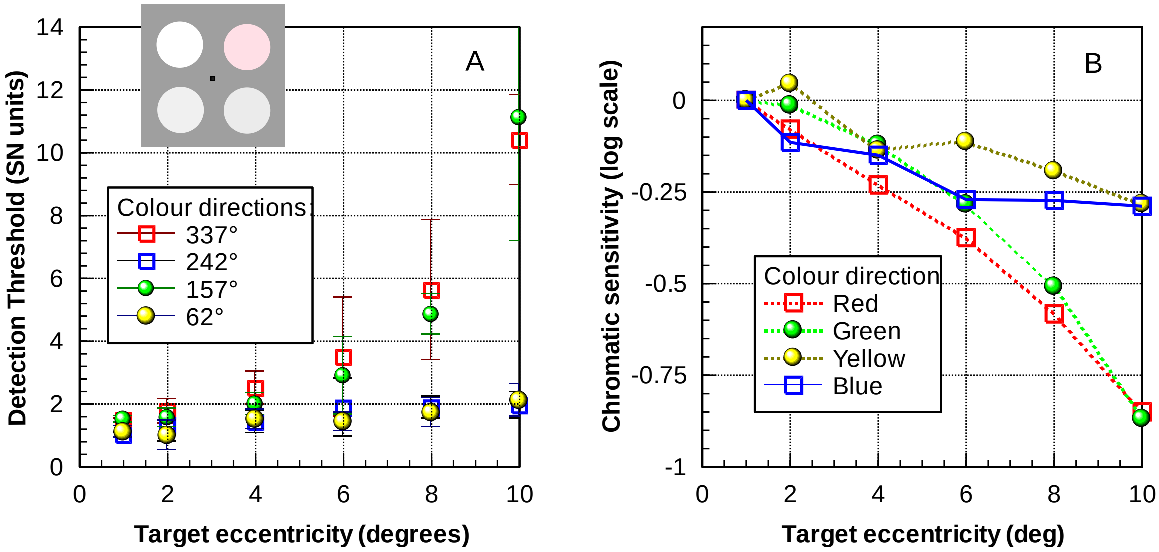

Figure 21: Variation in RG and YB colour vision with eccentricity. (A) Thresholds for detection of red, green, yellow and blue colours measured at a number of discrete eccentricities using a four-alternative, forced-choice procedure. Each data point shows the average threshold for 3 normal trichromats and the error bars indicate ±SD. For convenience the thresholds are expressed in CAD units. Each disc (see inset) subtended 25 min arc and its luminance contrast was selected randomly in the range 50% to 70% to avoid the use of luminance contrast cues when a colour signal is added to one of the four discs (see inset). The subject's task was to press one of four buttons to indicate the location of the coloured disc. (B) Chromatic sensitivity, i.e., the reciprocal of the threshold colour signal (plotted on a log scale) for each of the four colour directions employed.

In summary, YB colour signals have some advantage over RG, largely because YB chromatic sensitivity falls off less rapidly with increasing distance on the retina between the point of regard and the target location. RG colour signals, particularly when small targets are involved, have advantages over YB signals in central vision.

TCTs in subjects with colour vision deficiency

When no spatial feature is used to define the target (see Fig. 16 D, E&F), the task is more demanding and the subjects need to see and register correctly the colour of the target in order to locate and identify it. Data blocks contain detailed information, however this information must be read in a serial manner in order to determine which are relevant and should be 'attended' and which are not. There will therefore often be no guiding spatial cue that delineates target data blocks from other information on the display. Efficient and effortless processing of visual information and the ability to sustain a high level of performance over long periods of time benefits greatly from the use of colour signals. Multiple coloured targets are often presented on ATC displays and the use of colour makes segmentation and identification of objects into different categories easy to achieve. This in turn leads to more efficient visual search and overall improvement in visual performance.

The use of multiple colours over a large visual field and the absence of additional spatial cues to define the test target is arguably the most demanding use of colour in visual displays. This corresponds to Class IIIB experiment as described in section 3.3.4. It was therefore decided that this type of test should be carried out with a larger number of normal trichromats and subjects with varying degrees of colour vision loss. Multiple coloured distractors were shown on the screen and both target and distractors were Landolt Cs. The target could therefore only be located if the subject had adequate chromatic sensitivity to do so. The colour hues that isolate RG and YB chromatic mechanisms can be more challenging when the subject has reduced chromatic sensitivity as a result of congenital deficiency. With respect to the background employed in CRATO tests these colours correspond to chromatic displacement directions of: 337° (red), 157° (green), 62° (yellow) and 242° (blue). Only one of these colours was assigned to the target in each experiment and the remaining three colours and the many achromatic Landolt Cs acted as distractors. The chromatic saturation of the target was set at 12 CAD units. This suprathreshold colour signal is easily achievable on visual displays. Subjects with normal trichromatic colour vision also show little or no improvement in TCTs for chromatic displacements above 12 CAD units (see Fig. 18). In other words, TCTs reach an asymptote well before 12 units. Luminance contrasts of targets and distractors were set at ±45%, values that are well above detection thresholds for the target sizes employed.

Figure 22 shows visual search times and performance scores (expressed as percentage correct responses), for normal and colour deficient subjects in Class IIIB experiments. Fifteen normal trichromats and ten deutan subjects completed this experiment. The results show that on average colour deficient subjects cannot carry out the visual search task as quickly as normal trichromats and consequently they end up with significantly longer TCTs. Deutan subjects with minimal deficiencies, i.e. RG colour thresholds below 3 CAD units, can however carry out the task accurately with similar percent correct scores. The negative luminance contrast condition tends to produce longer TCTs for both red (Fig. 22A) and green (Fig. 22B) targets, within the deutan group. In contrast, for both yellow (Fig. 22C) and blue (Fig. 22D) target colours, and for both luminance contrast conditions, the colour deficient subjects perform as well as normal trichromats both in terms of speed and response accuracy. It is worth noting that not all subjects achieved 100% correct responses and this may well be due to errors in using the keypad, or other lapses in attention, rather than an inability to register accurately the colour of the target. This observation applies both to normal trichromats and to colour deficients. Only results for normal trichromats and deutan subjects are presented since only minimally affected deutans can pass Class IIIB experiments. Several protans were also tested, but everyone failed to complete the test for RG colours.

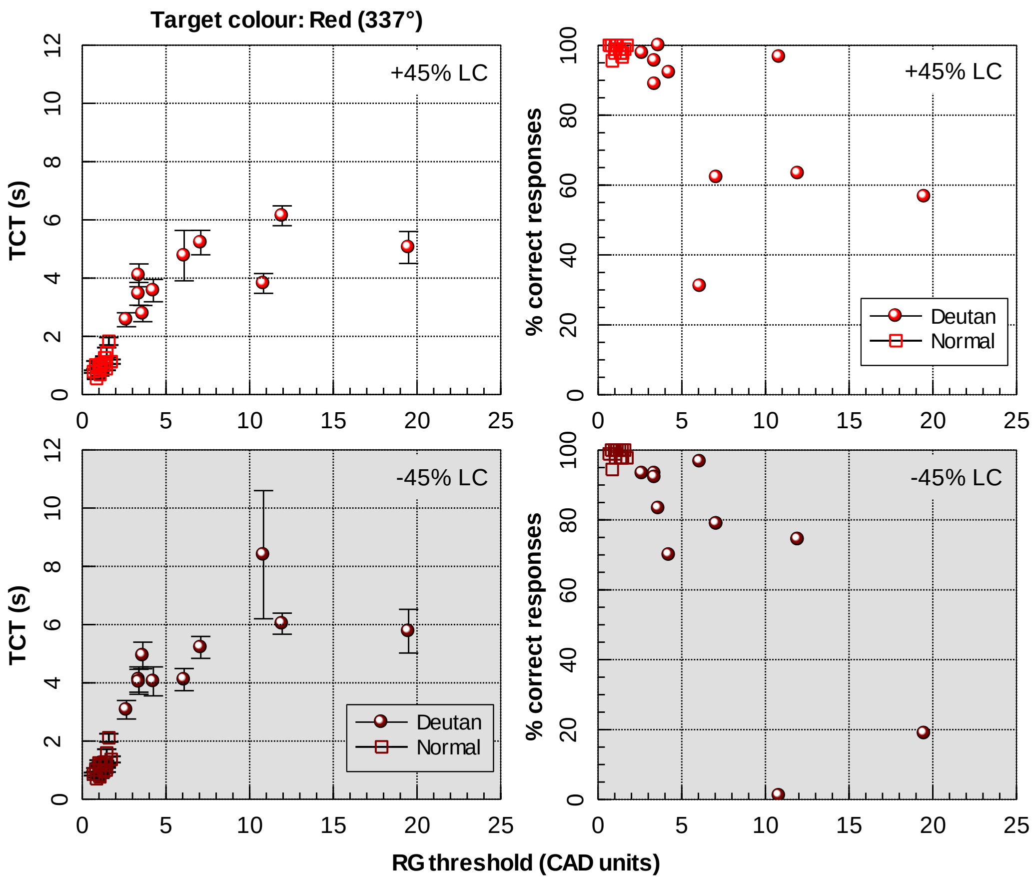

Figure 22a: Results of Class IIIB experiments in normal trichromats and colour deficient subjects. TCTs and CRS values plotted against the subject's RG CAD threshold. Both targets and distractors were Landolt Cs, 40 distractors (3 coloured). The chromatic signal strength was set at 12 CAD units. Luminance contrast (LC): +45% and -45%; colours displayed: 64o, 157o, 244o and 337°; colour investigated: red (337°).

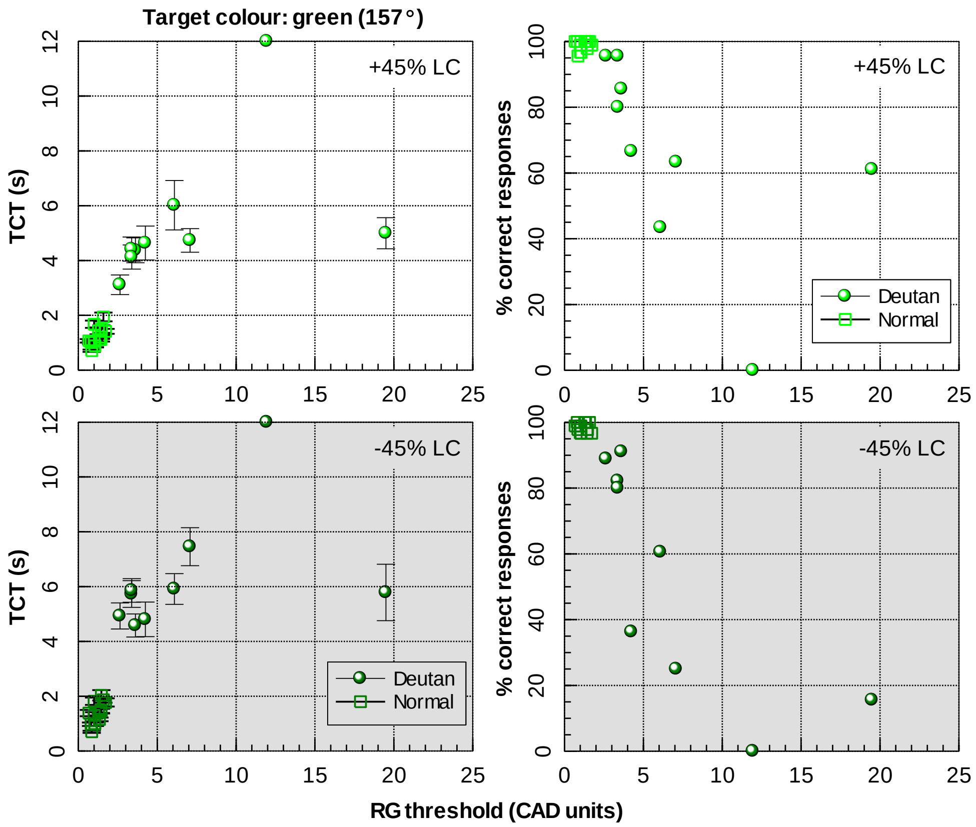

Figure 22b: Same as Fig. 22A, but the colour investigated was green (157°).

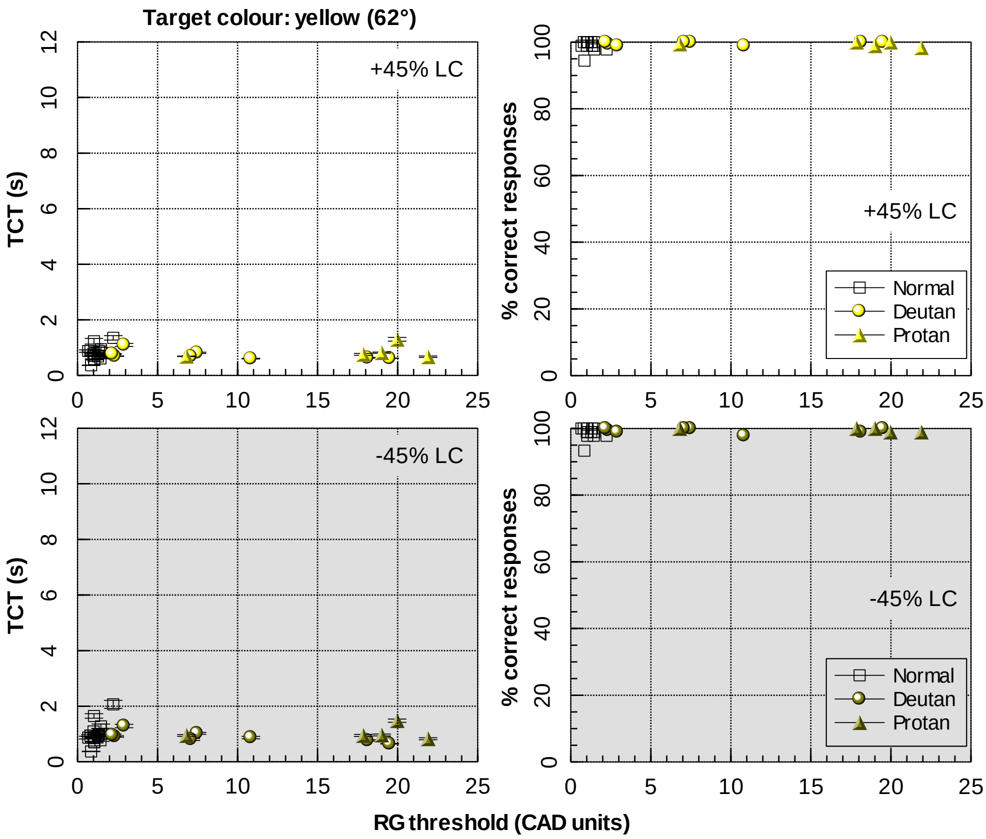

Figure 22c: Same as Fig. 22A, but the colour investigated was yellow (62°).

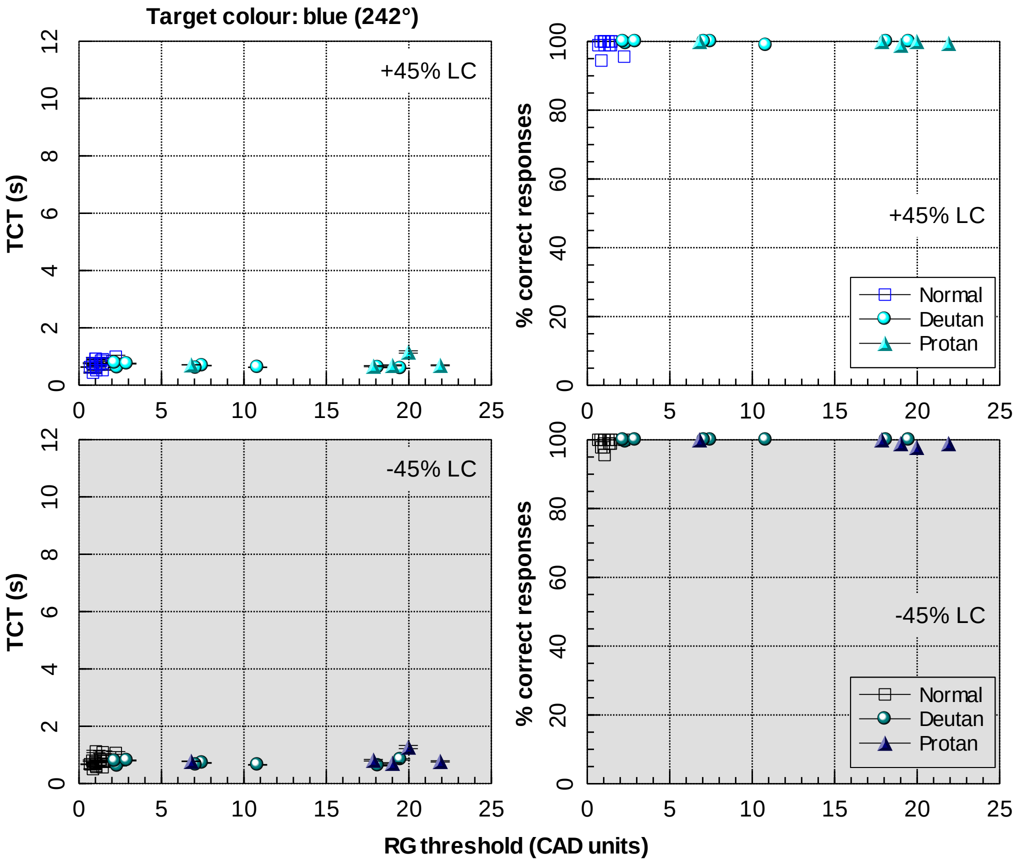

Figure 22d: Same as Fig. 22A, but the colour investigated was blue (242°).

Pastel colours – targets with suprathreshold RG and YB components

The experimental findings so far have shown that for yellow and blue targets congenital colour deficient subjects can perform, in terms of TCTs and correct scores, in a similar way to normal trichromats. In contrast, for the red and green targets, only the mildest colour deficient subjects managed to achieve the percentage correct scores of normal trichromats and in general, they all produced longer TCTs. Since colour hues and chromatic saturations can technically be altered in ATC applications to suit individual operators, it is of interest to examine whether suprathreshold colours with both RG and YB colour differences (that are conveniently although not always correctly described as pastel colours) can be used by subjects with congenital RG colour deficiency to achieve the levels of performance that are associated with normal trichromatic colour vision. To test this possibility, we repeated the Class IIIB experiments with four suprathreshold colours selected to ensure that each colour generates both RG and YB colour differences.

The aim of this experiment was therefore to establish whether use of pastel colours, as defined in Figure 23, yield visual performance equivalent to normal trichromats.

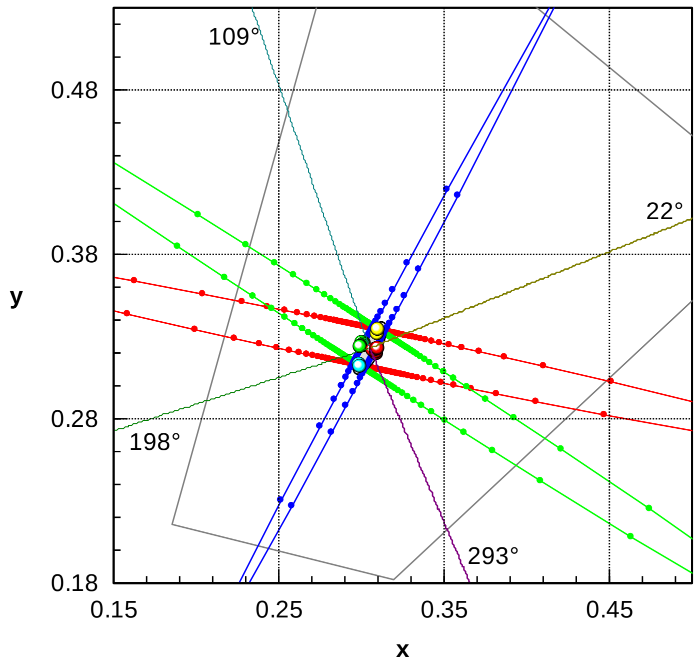

Figure 23: New colour directions (lines at 22°, 109°, 198° and 293°, plotted in CIE 1931 chromaticity chart). The central point indicates the background chromaticity. Note that for the background chromaticity shown, the colour confusion lines correspond to 64° and 244° for YB discrimination, and 337° and 157° for RG discrimination. Protan, deutan and tritan confusion axes with respect to the selected background field are indicated by the corresponding red, green and blue dots, respectively. The phosphor limits of the monitor are indicated by the grey lines.

For simplicity, the background chromaticity remained unchanged, but new colour directions of 22°, 109°, 198° and 293° (as shown in the in CIE-x,y 1931 chromaticity chart) were selected to ensure that each of the four potential coloured targets had a RG and a YB colour contrast component. Any other directions that are closer to the YB axis than those selected for this study (see Fig. 23) will generate even greater YB colour differences which RG colour deficients can use. Screen dumps showing coloured distractors drawn in the new (pastel colours) and the old (RG and YB axes isolating colours) are shown for comparison in Fig. 24).

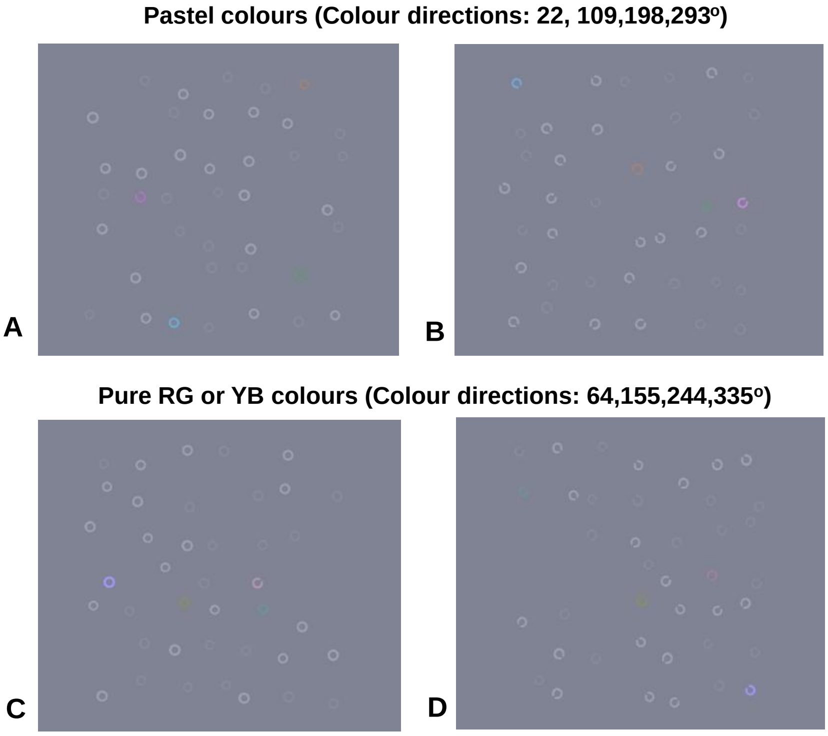

Figure 24: Sections A and C show stimuli for Class IIIA experiments when the target is defined by both colour and spatial cues. Section B and D display Class IIIB experiments when the subject has to search for the named target colour. 'Pastel' colours generated along chromatic displacement directions of 22°, 109°, 198° and 293° are shown in sections A and B. Colours generated along directions that isolate RG and YB mechanisms (i.e., 64°, 157°, 244° and 337°) are shown in sections C and D. Following the detection of a named colour (selected in advance), the subject has to report the orientation of its gap.

TCTs were measured in 31 normal trichromats and 39 congenital colour deficients (25 deutans and 14 protans) in Class IIIB experiments, for each of the four colour directions proposed to ensure the colours generated were suprathreshold for both RG and YB chromatic mechanisms. The subjects with congenital colour deficiency varied in the severity of RG colour loss from almost normal thresholds to complete absence of RG colour vision. Figure 25 shows the results for each of the four pastel colours investigated in normal trichromats and in deutan and protan subjects. TCTs and the corresponding correct scores are plotted as a function of subject's RG CAD thresholds.

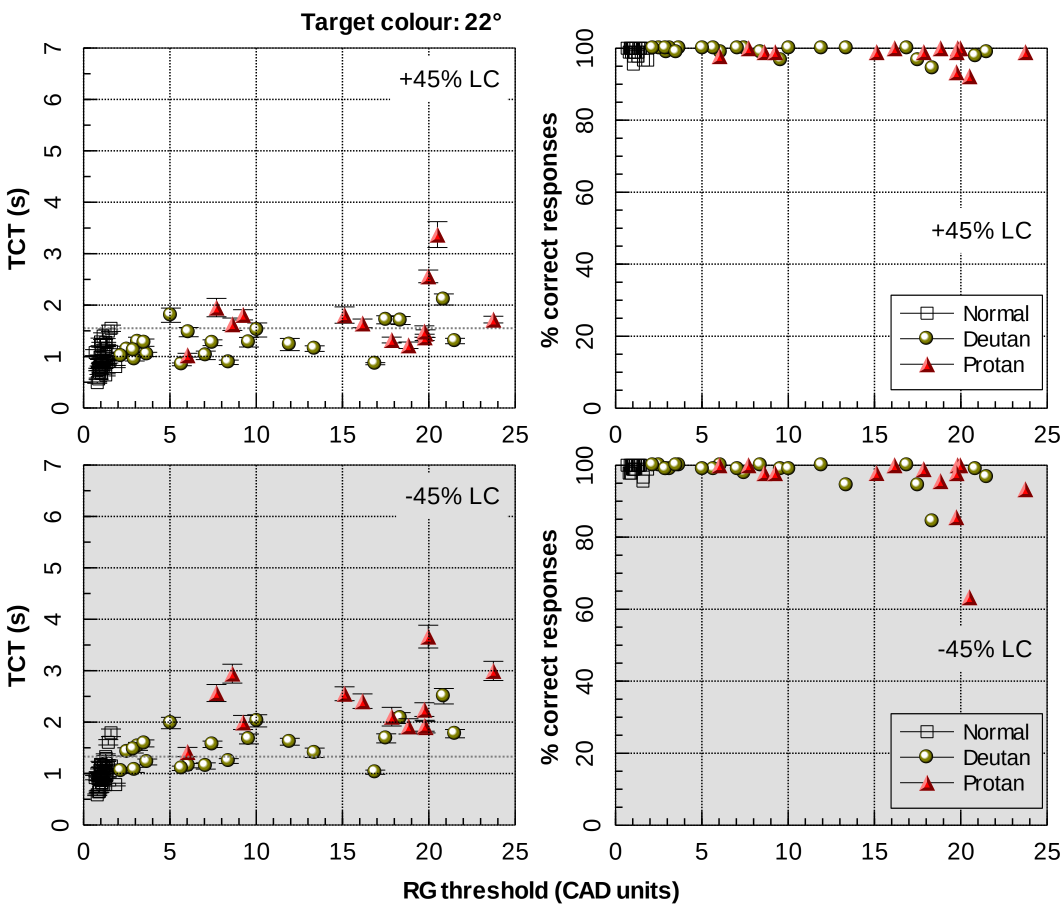

Figure 25a: Results of Class IIIB experiments with the modified pastel colours (generated along directions of 22°, 109°, 198° and 293°) in normal trichromats and colour deficient subjects. TCTs and CRS values plotted against the subject's RG CAD threshold. The results shown are for the 'orange' direction of 22o. Both targets and distractors were Landolt Cs, 40 distractors (3 coloured), chromatic signal strength: 12 CAD units. Luminance contrast (LC): +45% and -45%. Data are shown for 31 normal trichromats, 25 deutan and 14 protan observers.

The results for normal and colour deficients show that the addition of a YB colour signal to a target defined by luminance and RG colour contrast can benefit significantly both deutan and protan subjects. In fact, some of the colour deficients achieve TCTs and CRSs that are equivalent to those measured in normal trichromats. This was particularly the case for targets that also had positive luminance contrast. Although some deutans required slightly longer TCTs particularly for greenish (109°) and reddish (293°) colours when shown in negative luminance contrast, the correct response scores remained high and equivalent to normal trichromats even in deutans with large RG colour thresholds. Protans, on the other hand, take longer to carry out the visual search task for targets of both positive and negative luminance contrast, particularly those with thresholds above 7 CAD units. In spite of being on average slower to carry out the task, protans continue to produce accurate percent correct scores even when their RG CAD thresholds are as high as 20 SN CAD units.

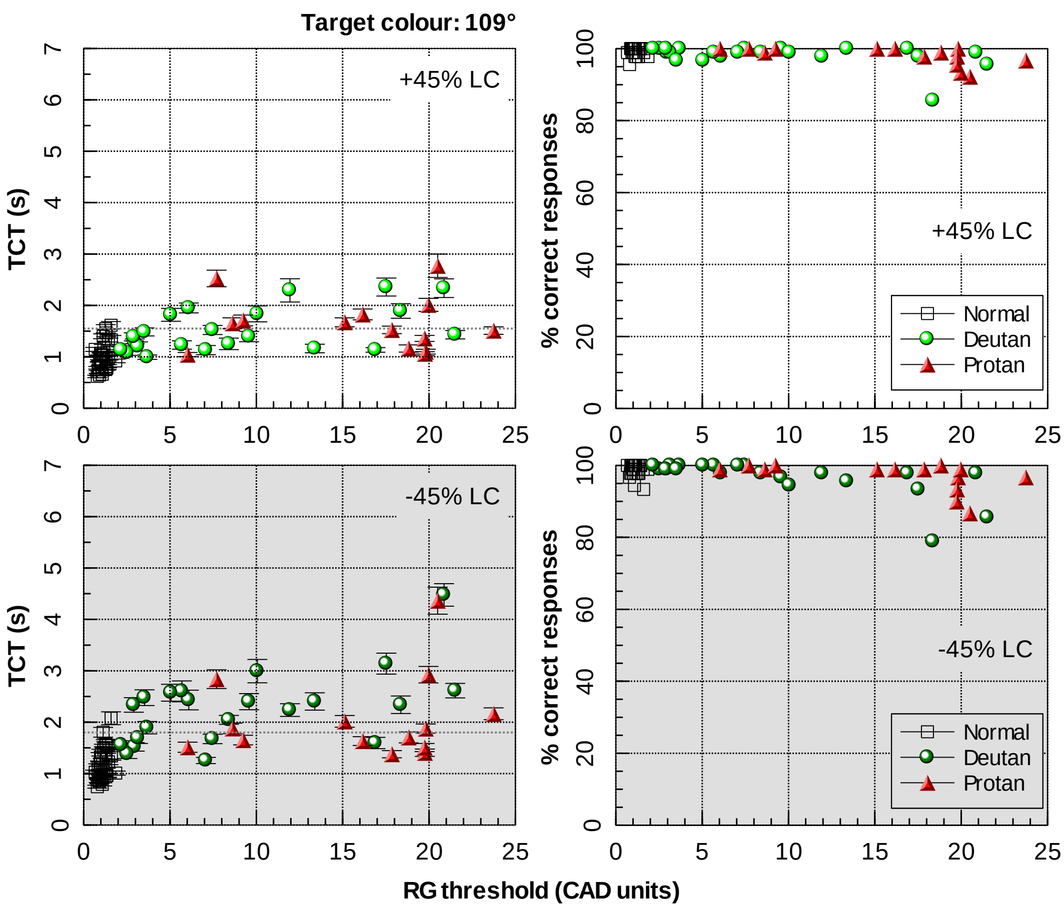

Figure 25b: Same as Fig. 25a, but the colour investigated was greenish (109°).

These modified colour directions improve visual search in subjects with colour deficiency. In addition, these directions do not impair the performance achieved in normal trichromats. In fact 'pastel' colours often produced shorter TCTs in normal trichromats when compared to the primary 'green' colour (i.e., 157°).

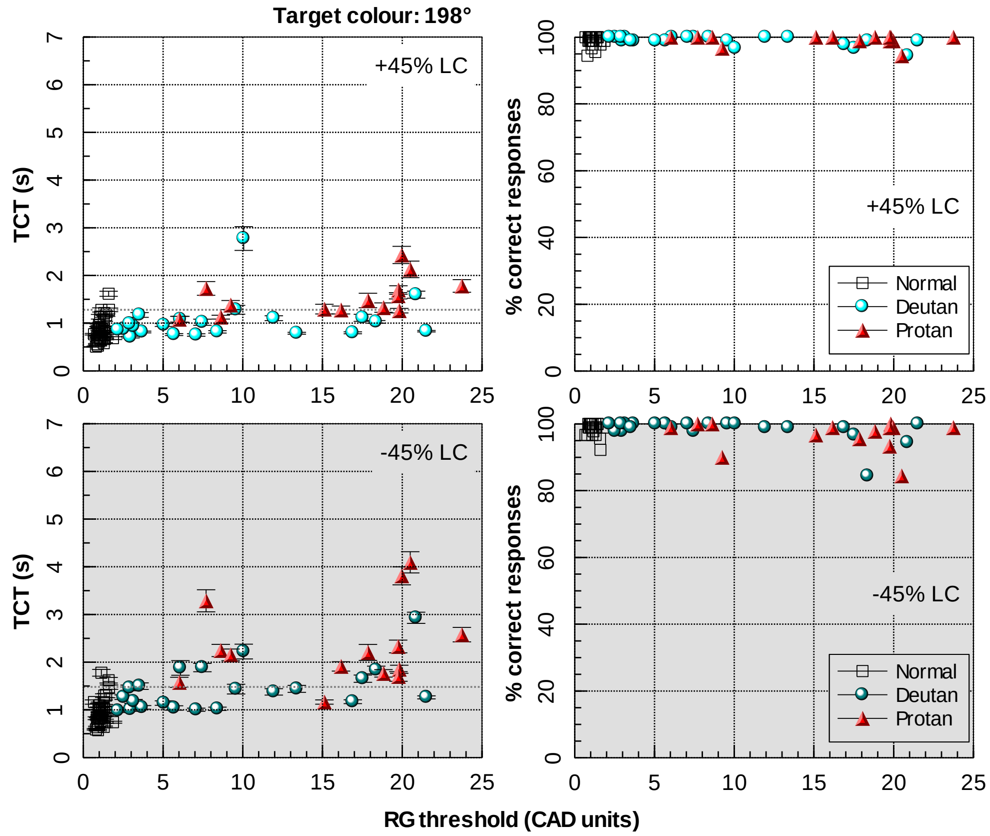

Figure 25c: Same as Fig. 25A, but the colour investigated was bluish (corresponding to an angle of 198°).

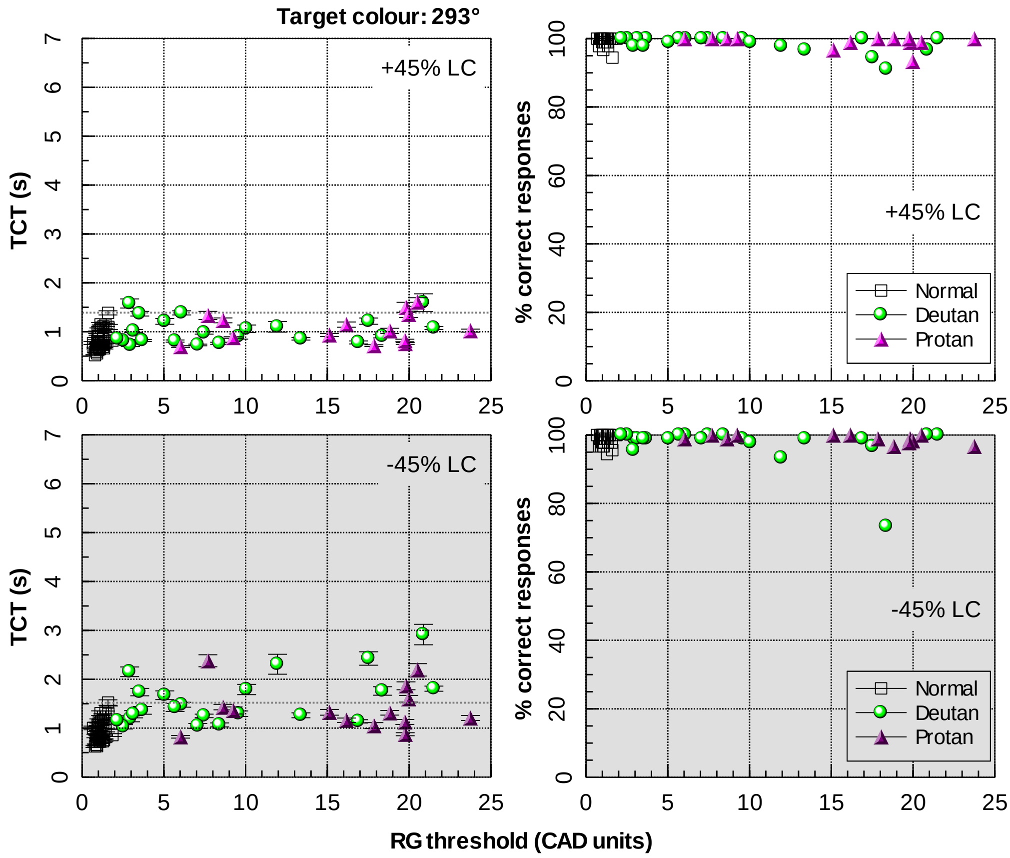

Figure 25d: Same as Fig. 25A, but the colour investigated was reddish (corresponding to an angle of 293°).

Demonstration of 'normal' performance with pastel colours in deutans

As part of a separate study designed to investigate eye-tracking and visual search we have examined how well a deutan subject with moderate deficiency (RG threshold of 4.8 CAD units) performs the visual search task when presented with multiple colours that isolate either RG or YB chromatic mechanisms (i.e., 64°, 157°, 244° and 337°). The chromatic saturation employed in this study was 10 CAD units above threshold for each of the four colours. Although not part of this study, the results are of interest and relevant to this report. The same experiments were then repeated with the four pastel colours, each of which generates both RG and YB chromatic signals (i.e., chromatic directions of 22°, 109°, 198° and 293°). In addition, the colour deficient subject carried out the Class I experiment with pure achromatic targets of 15% and 60% luminance contrast. Two normal trichromats also carried out the same tests to provide comparison data. The results are presented in Figures 26-28 and the key findings are summarised below:

- When the target is purely achromatic and defined by a spatial cue or when YB colour signals are added to the target, the deutan subject performs as well as normal trichromats. This is also the case when no spatial cue is available and the subject has to rely totally on the named colour of the target.
- The deutan subject produces 100% correct scores in all these tests.
- When the same experiments are carried out with suprathreshold colours that isolate the RG chromatic mechanism (i.e., 157° and 337°), the subject continues to produce 100% correct scores (see Fig. 26), but his TCTs are significantly longer than those measured with YB colours (Fig. 26A).
- When only the four pastel colours are employed in the visual search experiment (i.e., 22°, 109°, 198° and 293°), the subject's TCTs are similar for all colours and completely within the normal range. His scores are 100% correct for each colour in both Class IIIA (colour and spatial cue) and IIIB (colour only) experiment.

These findings confirm that the deutan subject is able to use each of the four pastel colours to find the target of interest with the same speed and accuracy as normal trichromats.

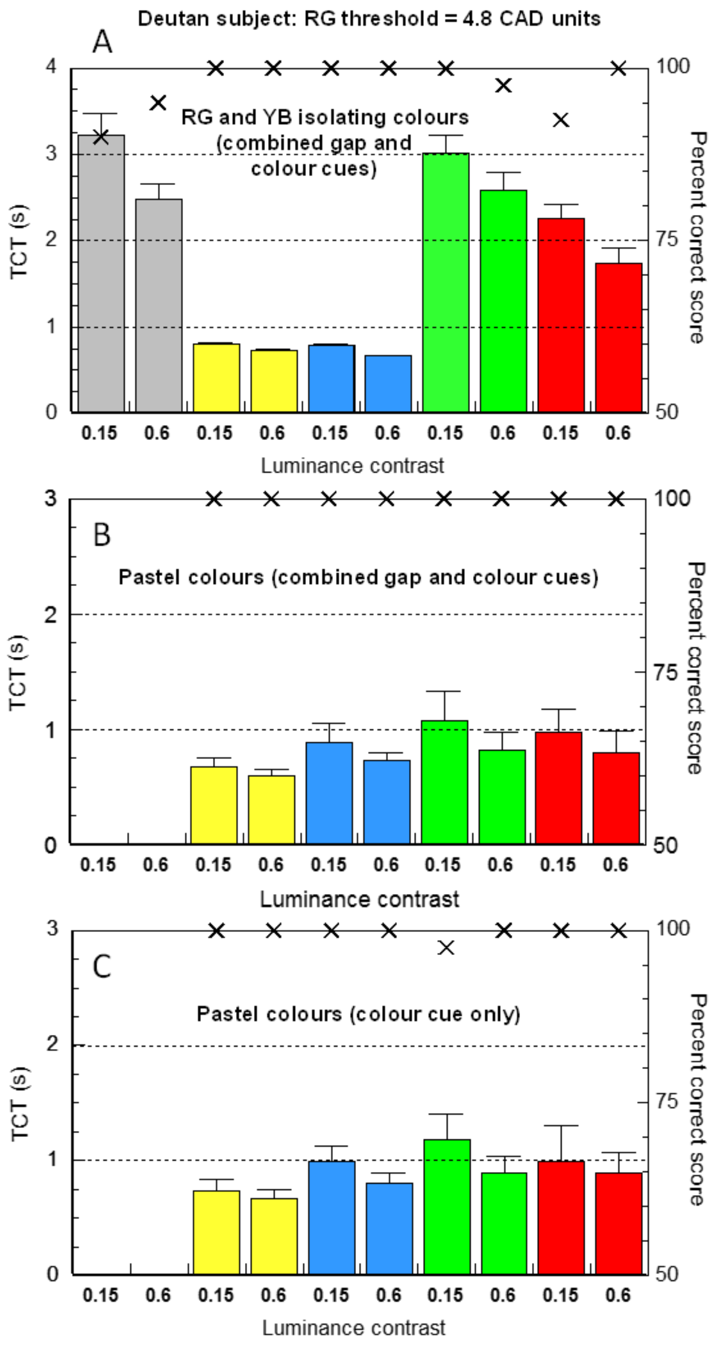

Figure 26: Results for Class IIA (A), IIIA (B) and IIIB (C) experiment for a deutan subject with a RG threshold of 4.8 CAD units. Coloured bars show TCTs for a chromatic saturation of 10 CAD units for, from left to right, yellow, blue, green and red. Targets and distractors had a percentage luminance contrast of 15% and 60% with respect to the uniform background. The grey bars show results for the achromatic condition (Class I experiment) for 15% and 60 % luminance contrast. The results show that this deutan subject performs as well as normal trichromats (see Figs. 27 and 28) when the target is purely achromatic and defined by a spatial cue or when YB colour signals are added to the target. The subject's TCTs are similar for all colours and completely within the normal range when pastel colours are employed (B and C). When RG isolating colours are employed the subject requires longer TCTs compared to YB isolating colours. Percent correct scores are also displayed (black crosses) and reveal practically 100% scores on all the tests.

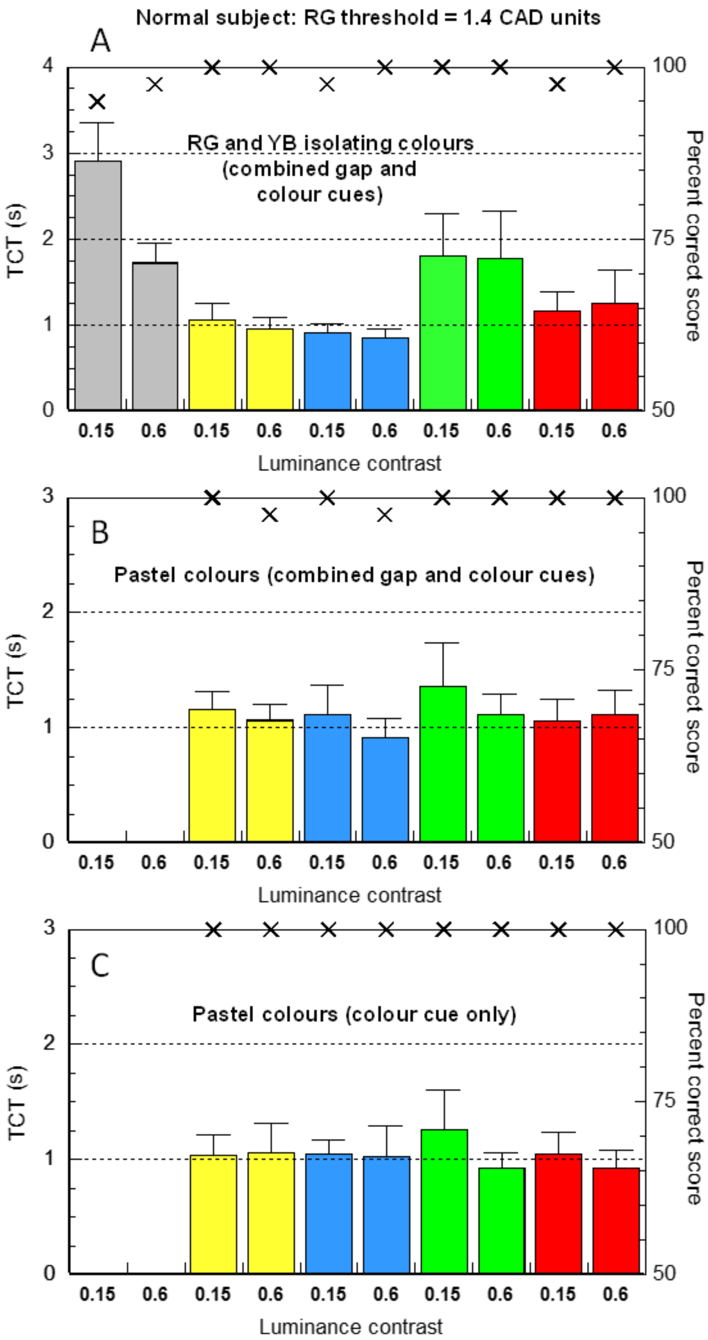

Figure 27: Same as Fig. 26, for a normal trichromat with a RG threshold of 1.4 CAD units.

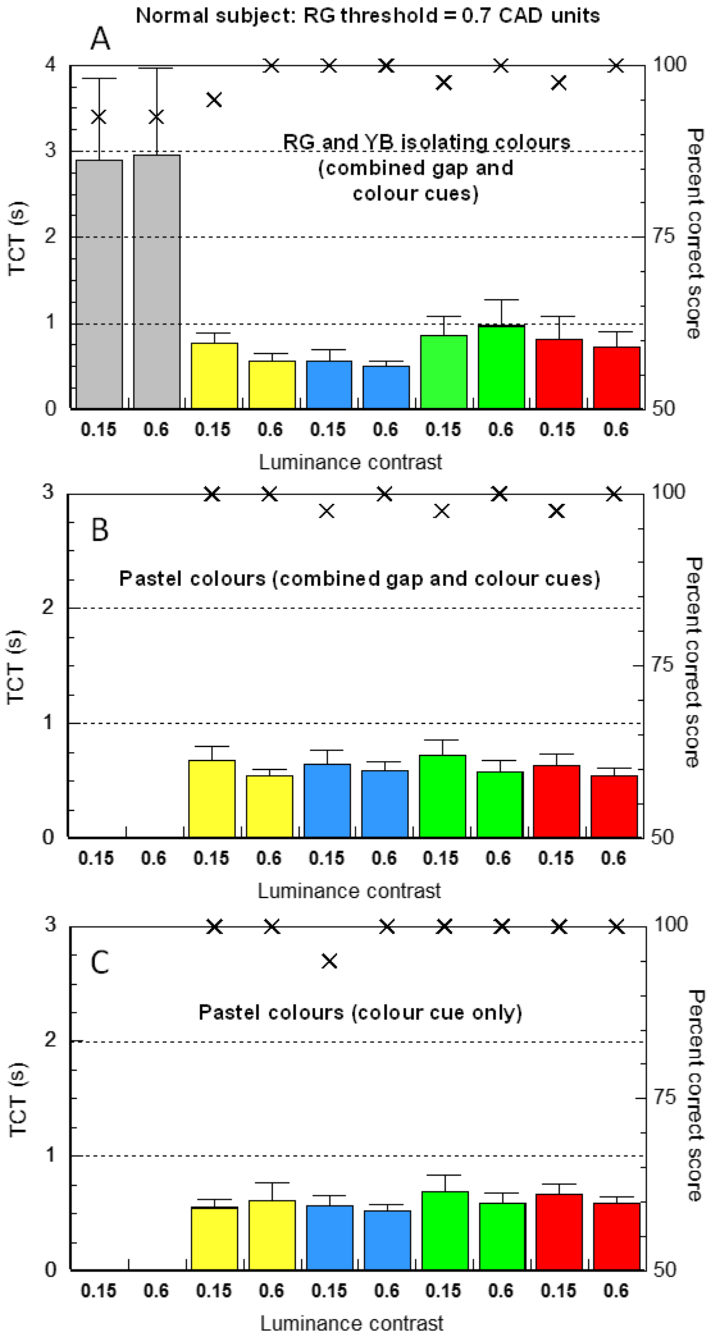

Figure 28: Same as Fig. 26, for a normal trichromat with a RG threshold of 0.7 CAD units.

Chapter 4

Analysis of HW-A

4.1 Data and correlation with severity of RG loss

The ability to name correctly the colours of signal lights is important in some ATC applications, particularly when work in ATC tower centres is involved. Performance on the HW-A lantern as recommended by the CIE and used by the UK CAA has provided for decades confirmatory or secondary testing to assess whether candidates with congenital colour deficiency have sufficient chromatic sensitivity to carry out colour-related tasks within selected environments. The HW-A lantern is currently used by the British Armed Forces and prior to 2009 was the accepted secondary test for the UK CAA. HW-A is a useful lantern test which all subjects with normal trichromatic colour vision pass. The test employs red, green and white lights of small angular subtense and varying intensities (see section 2.3). The task requires the subject to name correctly the colours of the lights presented to the eye according to specific protocols. Only normal trichromats and mild deutan subjects pass these protocols (CIE 2001; CAA 2009). No protans pass the HW-A lantern using the same protocols. With increasing severity of colour vision loss, deutan subjects exhibit greater variability in the number of errors they make in repeated tests. This remains a problem since on some occasions, an applicant may fail a first assessment, yet pass on repeated testing. We examined this problem by analysing statistically the errors subjects make and the pass / fail rates as a function of the severity of their colour loss. These results are presented in this report and the findings contribute significantly to the final recommendations.

Another reason for examining the HW-A lantern and for attempting to find equivalent methods of assessment that could replace the HW-A lantern is simply because this lantern is no longer manufactured or supported in terms of calibration and servicing. It is now known from previous studies that neither the Ishihara test nor the Nagel anomaloscope can be used to screen with certainty for normal trichromacy or to quantify the severity of colour vision loss in subjects with colour deficiency (Barbur et al. 2008; Rodriguez-Carmona et al. 2012). A large number of normal trichromats (~

10%) make at least one error on the Ishihara test (first 15 plates on the 24-plate version) and some subjects with minimal colour deficiency can pass the anomaloscope test. It is also worth noting that ~ 1.5% of deutans pass the 24-plate system with zero errors, but all protans fail. The more thorough examination using the first 25 plates of the Ishihara 38-plate edition results in improved sensitivity. When a pass requires zero errors, only ~ 0.6% of deutans pass and all protans fail. The disadvantage of using the 38-plate edition is that 19% of normal trichromats also fail when the pass criterion is based on zero errors. The UK CAA has always used the Ishihara test for the screening of colour vision and has used a secondary colour vision test (such as the HW-A which all normal trichromats pass) to determine ‘safe’ colour vision for a particular occupation. In order to obtain reliable data on pass / fail outcomes, we examined over 300 colour defective subjects (226 deutans and 92 protans) and 41 normal trichromats on the HW-A, Ishihara and the Colour Assessment and Diagnosis (CAD) test. Following the UK CAA protocol (outlined in Section 2.3), the Ishihara / HW-A passes ~ 22 % deutans, 1.1% protans and 100% of normals. The ~ 22% of deutans that pass the HW-A lantern and Ishihara test using the pre-2009, UK CAA protocol (based on two or less errors on the first 15 plates of the 24 plate Ishihara test followed by the HW-A lantern) were indistinguishable from normal trichromats (see Table 2). Although these deutans are in general those with least severe loss of RG colour vision, more moderate deutans can also pass. The latter group tend to exhibit greater variability with increased number of mean errors on repeated runs. The data presented in Figure 29 are of great interest since they show that even those deutans that pass can make colour naming errors on the HW-A lantern test on repeated runs. The results are based on six runs carried out in each of the 226 deutan subjects. Although subjects in both groups make a number of mean errors per run, those that fail make significantly more errors per run than those that pass. For example, the least affected 50% of those that pass make 0.6 or less errors per run. The equivalent number for those that fail is 3.4 or less errors per run. Although statistically these observations are not surprising, it is worth noting that because of the large crossover, some of the subjects that fail the HW-A test make less errors per run than some of the subjects that pass.

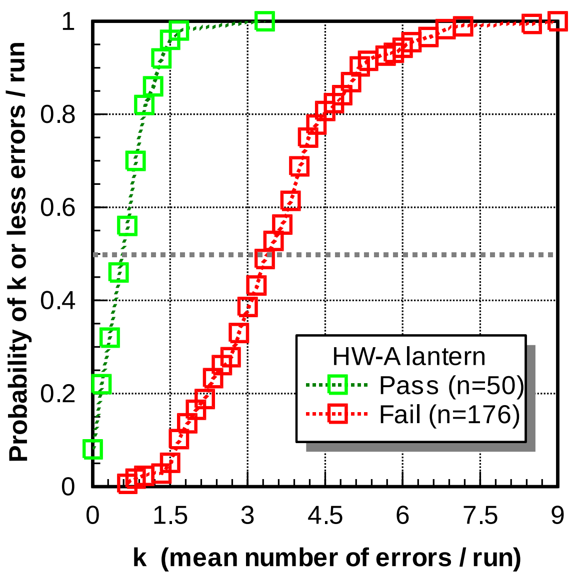

Figure 29: Plot showing the probability of making k or less errors per run for the deutan subjects that pass and for those that fail the HW-A protocol.

The cumulative frequency distributions show that 50% of those that pass make less than 0.6 errors / run, but those that fail make less than 3.4 errors / run. Each subject carried out a total of 6 runs (i.e., 54 presentations of the paired vertical lights).

The results show that several of the subjects that failed the standard protocol make fewer errors than many of the subjects that pass. This does mean that some deutans that pass in some runs may fail the same procedure in repeated runs. Deutans that passed and failed the HW-A lantern are indicated by green and red squares, respectively.

As mentioned in Section 1.3, the current requirements for European Class 3 medical certification is normal trichromatic colour vision as stipulated by the Ishihara test and the CAD test. This means that since 2009 when this approach was implemented ~22% of deuteranomalous applicants that have in the past, been classed as safe, no longer pass. The important observation that emerges from this analysis (which accounts for the ranked data shown in Fig. 30) is that deutans with intermediate levels of RG colour loss exhibit greater variability in their scores. They can therefore either pass or fail in repeated runs.

Figure 30 shows the distribution of CAD RG thresholds of deutans that pass and fail the HW-A lantern according to this protocol. Not unexpectedly, the results reveal that there is an overlap in RG thresholds measured on the CAD test among deutans that pass and fail this lantern. This is caused largely by the inherent within subject variability on repeated measurements on the HW-A lantern and the CAD tests (Squire et al. 2005). Although the mean number of errors per run on the HW-A test would provide a less variable pass / fail criterion, this would require several repeats and is therefore impractical.

The findings in Figure 30 also show that all deutans with a threshold less than 2.35 RG CAD units (~ 6%) pass the HW-A. A useful analysis of the overlapping section of deutans that pass and fail (lower part of Fig. 30), reveals that a RG threshold value of 4 units ensures an equal number of false positives and false negatives. In other words, the number of deutans that pass the standard protocol with a threshold > 4 units equals the number of deutans that fail with a threshold ≤ 4. The RG CAD threshold of 4 units passes 22% of deutans which is equivalent to the pass / fail outcome reported previously (CAA, pre-2009 protocol).

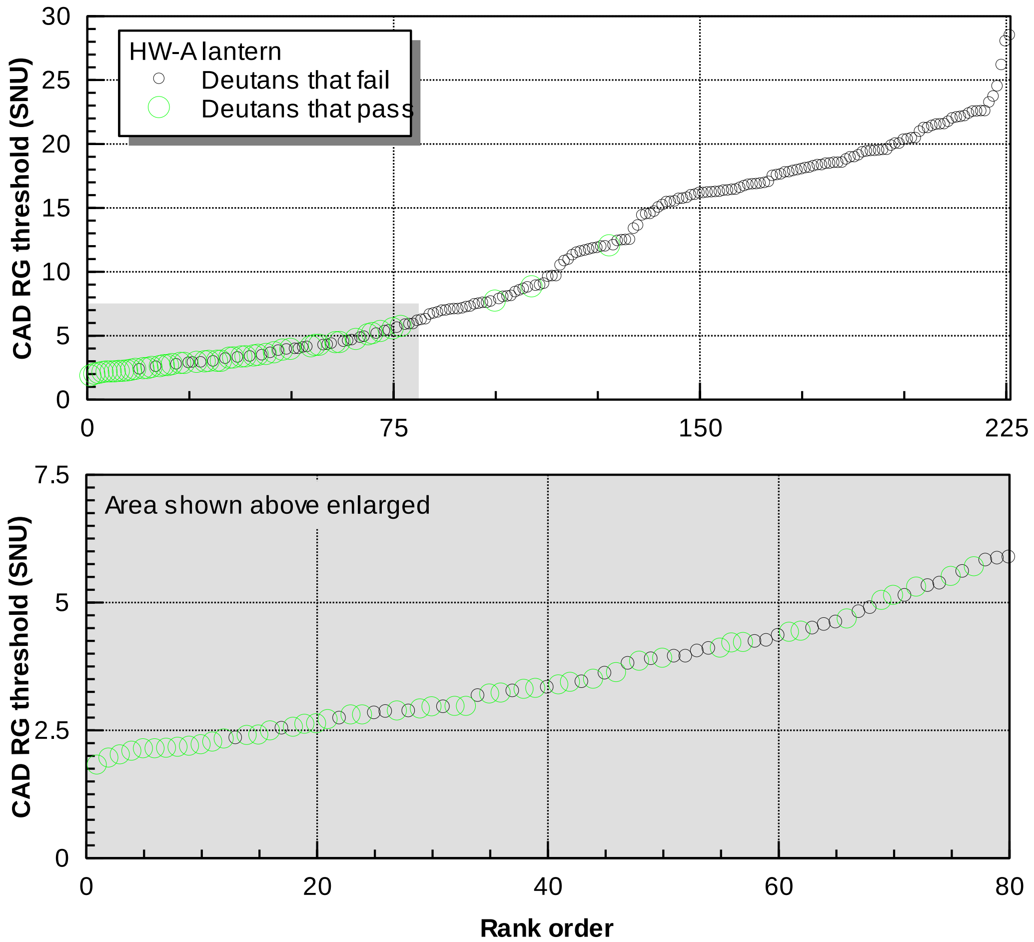

Figure 30: Ranked distribution of CAD RG thresholds for the 226 deutan subjects examined in the study. Deutan subjects that failed or passed the HW-A lantern according to the CAA protocol (Section 2.3) are indicated by black or green circles, respectively. Section B shows that shaded area of section A on an expanded scale for clarity of presentation. All deutan subjects with RG CAD thresholds < 2.35 units pass the HW-A lantern with close to zero errors.

4.2 Summary of outcomes based on HW-A lantern scores

By examining the HW-A scores in relation to the subject's severity of RG colour vision loss, CAD threshold limits can be recommended that are statistically consistent with the outcome of past practices. Although a value of 4 SN CAD units is statistically equivalent to the outcome of past practices, it is worth noting that applicants with even higher levels of deutan deficiency have passed a HW-A lantern in the past and are currently working safely in the ATC environment with no known problems. This demonstrated ability supports the argument that in the absence of a detailed study of colour vision requirements within the ATC environment, the recommended limit of 4 SN CAD units is safe and hence justifiable.

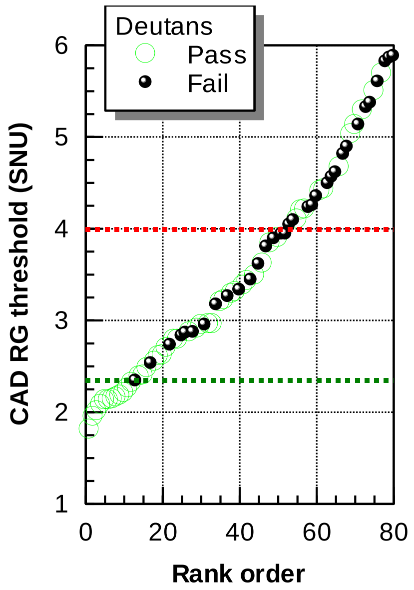

Figure 31: Magnified section of Fig. 30 with the proposed limits of 2.35 and 4 units. When these limits are applied, respectively, ~ 6% and 22% of deutans pass. No protans pass since with a few exceptions their thresholds tend to be above 5 CAD units.

The results of Fig. 30 and the evidence gained from past practice using protocols based on Ishihara and HW-A lantern suggest that two limits emerge that can be used to separate mild congenital colour deficients into two categories. These limits ensure that subjects perform well when the task involves discrimination and naming of red, green and white lights. The new categories can be summarised as follows:

- Subjects with thresholds ≤ 2.35 can be classed as safe and virtually equivalent to normal trichromats. In general, these subjects do not make errors on repeated HW-A tests.
- Subjects with thresholds ≤ 4 CAD units can perform all CRATO tests that employ suprathreshold ‘pastel’ colours with the same speed and accuracy as normal trichromats. These subjects can also discriminate red, green and white lights, although the statistical distribution expected for those with a mean threshold close to four units means that some of these subjects may pass and then fail the standard protocols on repeated tests. This point is illustrated clearly in Figure 29 which shows that some deutans that fail the HW-A make fewer mean errors per run than the deutans with the largest RG thresholds that pass the same protocol. The two RG CAD limits of 2.35 and 4 are illustrated in Figure 31 together with the subjects that fail the standard protocol with a RG threshold < 4 CAD units and those that pass with threshold > 4 units.

Chapter 5

Discussion

5.1 Colour vision concerns in ATC

Extensive research studies have been carried out by the FAA and other organisations into the effectiveness of using colour in ATC displays. Several human factors guidelines have been produced to advise on the use of colour in ATC applications. The progress made in colour display applications has been very rapid and in spite of what has been achieved; much remains to be understood in order to produce national and international standardisation for the use of colour in ATC and for specifying minimum colour vision requirements for ATC controllers. The relative ease of customising colours to suit different users creates new opportunities and potential problems. The findings of this study can be used to maximise the benefits of using colour and also to enable some subjects with congenital colour deficiency to work in the ATC environment. In addition, the data may also be useful for display designers to optimise the use of colour / luminance combinations and to avoid potential colour related problems.

The results of CRATO tests are important for a number of reasons:

- The experiments yield TCTs and CRSs when single or multiple colours are used with targets similar in size to data blocks employed in ATC applications.
- The tests also yield equivalent visual performance data as a function of luminance contrast which provide a scale against which the advantages of colour signals can be assessed.
- The experiments assess how RG and / or YB colour signals benefit visual performance in visual search tasks that involve the processing of visual information over large visual fields.
- The results obtained separate the effects of colour signals when added to targets of positive and negative luminance contrast. These observations

are important in normal trichromats and reveal significant differences in subjects with congenital deficiency.

- The study identifies suprathreshold use of colour and stimulus conditions when many subjects with congenital deficiency perform as well as normal trichromats.

In addition, the analysis of HW-A lantern scores and the comparison with the applicant's severity of RG colour vision loss help provide RG threshold limits that are statistically equivalent to the old UK CAA colour assessment protocol (the latter are still being used by the Maritime Coastguard Agency, the British Army and the UK Electrical Contractors Association).

In relation to ATC applications, the following recommendations ensure that the applicants that pass can make full use of colour within large-field visual displays with the speed and accuracy expected of normal trichromats. In addition, those that pass would also be able to discriminate the red, green and white signal lights when varying in intensity (as employed in the HW-A lantern test).

1. All applicants with RG and YB CAD thresholds below the upper normal limits that have been established for healthy aging (Barbur and Rodriguez-Carmona 2015) should pass by default since they have normal trichromatic colour vision and are likely to exhibit the best performance that can be achieved. According to the new classification system these applicants would be classed as having **'Normal' trichromatic colour vision (CV1)**. This category is appropriate without any further considerations of changes to display requirements or signal lights.

2. All applicants with a RG threshold ≤ 2.35 CAD units. These subjects exhibit almost normal RG colour discrimination and **pass the HW-A lantern test with zero errors**. In terms of anomaloscope match parameters the deutans that pass exhibit match ranges within normal limits, but require more 'green' in the red / green mixture field to match the monochromatic yellow field. These subjects are not likely to have any colour detection and discrimination problems when **suprathreshold colours** defined by both RG and YB components are employed in visual displays. No protan subjects can be included in this category. This limit is

sufficient to pass all normal trichromats, irrespective of age and ~ 7% of the least affected deutans. The latter exhibit almost normal RG colour discrimination and pass the HW-A lantern test protocol consistently in repeated tests. According to these observations and the results of CRATO tests, subjects with **'Functionally normal' trichromatic colour vision (CV2)** perform as well as normal trichromats when suprathreshold colours are involved. This category is appropriate providing the display requirements are designed to ensure that saturated, suprathreshold colours are employed, which is usually the normal practice.

3. Applicants with YB CAD thresholds within the normal range and RG thresholds ≤ 4 CAD units can also work as Air Traffic Controllers provided certain restrictions are imposed on the use of colour signals in ATC displays. The higher limit is sufficient to pass all normal trichromats and ~ 22% of deutan subjects. This higher limit matches the percentage of deutans that pass the HW-A lantern (22%) using the CIE protocol that has been recommended for use with this lantern. No protan subjects are included in this category since their minimum thresholds are in general larger than five SN units. Some of the deutans included in this group will have RG colour discrimination difficulties with small RG colour signals that are close to normal thresholds. These subjects do however exhibit normal levels of visual performance when suprathreshold colours defined by both RG and YB components are employed, which is often the normal practice in visual displays. This category can be described as **'Safe' trichromatic colour vision (CV3)**.

5.2 Summary and recommendations

The enhancement of key aspects of visual performance as a result of adding RG and / or YB colour signals in large field visual tasks has been investigated in normal trichromats and in subjects with congenital colour deficiency. Speed of performance and accuracy have been measured and related to the subject's RG and YB colour vision sensitivity. The recommendations put forward in this report are based largely on these findings and also the correlation between CAD thresholds and pass / fail error scores on Ishihara test plates and HW-A lantern tests measured in over 1000 subjects with both normal trichromacy and congenital colour deficiency.

In the absence of detailed studies designed to establish minimum colour vision requirements for specific occupational tasks (as has been done for flight crew), an acceptable alternative is to consider carefully the three categories described above and to select the one that can be considered safe, without discriminating unfairly against those subjects with congenital colour deficiencies that achieve levels of performance equivalent to normal trichromats.

- If the visual task requires detection and naming of colours for small signal lights (e.g. red, green, yellow, blue and white, etc.), or the discrimination of the smallest possible colour differences in order to judge uniformity of colour reproduction in manufactured goods, or the need to adhere to the commonest appreciation of perceived colour appearance and colour names and / or the ability to use efficiently faint, desaturated colours to segment objects into groups on visual displays, a **CV1** pass may be justified.
- When the visually-demanding, colour-related tasks involve the use of suprathreshold colours (e.g. with chromatic saturations and colour differences that are well above detection thresholds, which is the normal practice when colour is used deliberately to enhance visual performance in working environments), a CV2 pass can be allowed without compromising either efficiency or safety. Applicants with a CV2 pass can discriminate small RG colour differences, make very few errors on the Ishihara plates test, and more importantly, make no errors with signal lights that are equivalent to the HW-A lantern (i.e., small, often diffraction limited whites, greens and reds of varying intensity). Based on these findings, a **CV2** pass would be appropriate for air traffic controllers and seafarers (i.e., lookout officers).
- The CV3 category is appropriate for the majority of normal working environments that employ suprathreshold colours and do not involve discrimination of fine colour differences or the need to make correct judgements of colour appearance. Subjects with a CV3 pass will have

sufficient RG chromatic sensitivity to carry out suprathreshold, colour-related tasks, even when RG colour signals align along the corresponding colour confusion lines. CV3 is an important category for a number of reasons and hence benefits from further justification. Visual scenes in real working environments involve the use of large objects with spatial features that are often several times above the acuity limit. This is also the case when objects and images are generated on visual displays. The overall appearance of each object is determined by a combination of its size, luminance and YB and RG chromatic contrast. Differences in YB colour signals are often present in objects that are commonly classed as reddish and greenish and in red / white signal colours (such as the Precision Approach Pathway Indicator (PAPI) lights). If large chromatic saturations are employed, subjects with mild RG colour deficiency (e.g. those with a CV3 pass) will be able to make use of the reduced RG colour signal to carry out the colour-related task, even in those rare cases when YB colour differences are absent. The only slight disadvantage is that these subjects may be slower than normal trichromats when the tasks require visual search in large displays. One can, however, make use of this information and the novel findings that have emerged from this study in relation to the increased efficiency of YB signals in the periphery of the visual field to optimise the colours used in visual displays. When suprathreshold YB colour difference signals are also added to objects defined by luminance and RG colour contrast, congenital deficients that fall into the CV3 category can perform multi-colour visual search tasks with the same accuracy and speed as normal trichromats. Equally importantly, the majority of these subjects pass the HW-A lantern test, with equal number of false positives and false negatives centred with respect to the pass / fail limit of 4 CAD SN units. A **CV3** pass is therefore appropriate for applications that involve the use of large colour differences, particularly when these also involve YB colour signals. These findings show that with appropriate design and choice of colours for work in visual displays, the CV3 category can also be appropriate for use in the ATC environment as well as in many other occupations that involve the use of visual displays.

There is little doubt that subjects with congenital RG colour deficiency cannot perform as well as normal trichromats under the most demanding, unfavourable conditions when colour signals are close to colour detection threshold limits for normal trichromatic vision. This obvious conclusion does mean that many of these subjects cannot perform the most demanding, suprathreshold, colour-related tasks as efficiently and as accurately as normal trichromats. ATC applications involve well above threshold stimuli (defined by both colour and luminance contrast). Subjects with thresholds < 2.35 CAD can perform suprathreshold, colour-related tasks as well as normal trichromats. The findings described in this report also suggest that with appropriate design and choice of colours for work in visual displays, the CV3 category can also be appropriate for use in the ATC environment as well as in many other occupations that involve the use visual displays. If this category was adopted as appropriate, all normal trichromats and 22% of subjects with deutan deficiency would pass. Higher limits may also be appropriate for other specific occupations and working environments, but this can only be recommended with confidence if detailed studies designed to establish such limits are carried out.

Chapter 6

Acknowledgements

We wish to acknowledge the Civil Aviation Authority and the COLT foundation for supporting this work and City Occupational Ltd for programming the CRATO tests. We wish to thank in particular Dr Anthony Evans (ICAO) for the critical reading of the report and for their valuable comments and support throughout the project. We acknowledge the significant contribution each of the persons listed below has made to the project through progress review meetings and valuable criticism and comments on the final report:

Ms J. Birch	City University London, UK
Ms R. Dharmasiri	Surveillance Specialist, CAA, UK
Mr A. Harlow	City Occupational Ltd.
Mr M. Sutton	Air Traffic Control Officer, CAA, UK
Mr G. Greene	Research Projects Manager, CAA, UK

Finally, we wish to thank the many subjects that participated in the numerous experimental studies carried out as part of this project.

Chapter 7

References

Ahlstrom, V. & Longo, K. 2003, *Human Factors Design Standard (HF-STD-001)*, NJ: Federal Aviation Administration, Atlantic City International Airport.

Barbur, J. L. & Connolly, D. M. 2011. Effects of hypoxia on color vision with emphasis on the mesopic range. *Expert Rev.Ophthalmol.* 6, (4) 409-420.

Barbur, J. L. & Forsyth, P. M. 1988, "The effective contrast of coloured targets and its relation to visual search," *In Visual Search*, D. Brogan, ed., London: Taylor & Francis, pp. 319-328.

Barbur, J. L., Forsyth, P. M., & Wooding, D. S. 2003, "Eye movements and search performance," *In Visual Search 2: Proceedings Of The 2nd International Conference On Visual Search*, A. G. Gale, K. Carr, & D. Brogan, eds., Taylor & Francis, pp. 253-264.

Barbur, J.L., Harlow, A.J., & Plant, G.T. 1994. Insights into the different exploits of colour in the visual cortex. *Proc.R.Soc.Lond B Biol.Sci.*, 258, (1353) 327-334.

Barbur, J. L. & Rodriguez-Carmona, M. 2012, "Variability in normal and defective colour vision: consequences for occupational environments," *In Colour Design*, J. Best, ed., Elsevier Science, Woodhead Publishing.

Barbur, J. L. & Rodriguez-Carmona, M. 2015, "Color vision changes in normal aging," *In Handbook of Color Psychology*, A.J.Elliot, M.D.Fairchild, & A.Franklin, eds., Cambridge University Press.

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.*

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. *Vis.Neurosci.*, 25, (3) 507-516.

Barbur, J.L., Wolf, J., & Lennie, P. 1998. Visual processing levels revealed by response latencies to changes in different visual attributes. *Proc.R.Soc.Lond B Biol.Sci.*, 265, (1412) 2321-2325.

Belcher, S.J., Greenshields, K.W., & Wright, W.D. 1958. Colour vision survey using the Ishihara, Dvorine, Boström and Kugelberg, Boström, and American-Optical Hardy-Rand-Rittler tests. *Br J Ophthalmol*, 42, (6) 355-359.

Bergman, H. & Duijnhouwer, F. 1980. Recognition of VDU Presented Colors by Color Defective Observers. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 24, (1) 611-615.

Birch, J. 1997. Efficiency of the Ishihara test for identifying red-green colour deficiency. *Ophthalmic.Physiol.Opt.*, 17, (5) 403-408.

Birch, J., Barbur, J.L., & Harlow, A.J. 1992. New method based on random luminance masking for measuring isochromatic zones using high resolution colour displays. *Ophthalmic.Physiol.Opt.*, 12, 133-136.

Cardosi, K. & Hannon, D. 1999, *Guidelines for the use of colour on ATC displays*, U.S. Department of Transportation, Federal Aviation Administration, Washington D.C, DOT/FAA/AR-99/52.

Carroll, J., Neitz, J., & Neitz, M. 2002. Estimates of L:M cone ratio from ERG flicker photometry and genetics. *J.Vis.*, 2, (8) 531-542.

Carter, R.C. 1982. Visual search with color. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 127-136

Christ, R.E. 1975. Review and Analysis of Color Coding Research for Visual Displays. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 17, (6) 542-570.

Commission Internationale de l´Eclairage (CIE) 2001, *International recommendations for color vision requirements for transport.* International Commission on Illumination, CIE technical report, Vienna, Austria, Report no.143.

Civil Aviation Authority (CAA). 2014. *Air Traffic Services Safety Requirements, CAP 670* Airspace, ATM and Aerodromes, Safety and Airspace Regulation Group, Civil Aviation Authority.

Civil Aviation Authority (CAA). 2013, *European Class 3 (ATCO) Initial Examination Visual Standards*. Available online: http://www.caa.co.uk/Commercial-industry/Airspace/Air-traffic-control/Licences/Medical-requirements/European-Class-3-initial-examination-visual-standards/ (Accessed 22/7/2015).

Civil Aviation Authority (CAA) and Federal Aviation Administration (FAA). 2009, *Minimum Colour Vision Requirements for Professional Flight Crew - Part 3 Recommendations for a new colour vision standards*, The Stationary Office, Paper 2009/04.

Civil Aviation Authority (CAA). 2006. *Minimum colour vision requirement for professional Flight Crew - Part 1 The use of colour signals and the assessment of colour vision requirements in aviation*, The Stationery Office, Paper 2006/04.

Cole, B.L. & Macdonald, W.A. 1988. Defective colour vision can impede information acquisition form redundantly colour-Coded video displays. *Ophthalmic and Physiological Optics*, 8, (2) 198-210.

Cole, B.L., Maddocks, J.D., & Sharpe, K. 2004. Visual search and the conspicuity of coloured targets for colour vision normal and colour vision deficient observers. *Clin.Exp.Optom.*, 87, (4-5) 294-304.

EATM 2006, *Requirements for European Class 3 Medical Certification of Air Traffic Controllers.* Eurocontrol Headquaters, Brussels, HUM.ET2.ST08.10000-STD-02.

Frey, R.G. 1958. Most suitable pseudoisochromatic table for practice. Comparative studies on applicability of Boström, Böstrom-Kugelber, Hardy-Rand-Ritter, Ishihara, Rabkin and Stilling tables. *Albrecht.Von.Graefes Arch.Ophthalmol*, 160, (3) 301-320.

Holmes, J.G. & Wright, W.D. 1982. A new colour perception lantern. *Color research and application*, 7, 82-88

Kaiser, P.K. & Boynton, R.M. 1996. *Human Color Vision*, 2nd edition ed. Washington, D.C., Optical Society of America.

Kinnear, P.R. & Sahraie, A. 2002. New Farnsworth-Munsell 100 hue test norms of normal observers for each year of age 5-22 and for age decades 30-70. *British Journal of Ophthalmology*, 86, (12) 1408-1411.

Knoblauch, K., Vital-Durand, F., & Barbur, J.L. 2001. Variation of chromatic sensitivity across the life span. *Vision Res.*, 41, (1) 23-36 available from: PM:11163613

Mahon, L.E. & Jacobs, R.J. 1991. Electronic Flight Information System displays and colour defective observers. *Clinical and Experimental Optometry*, 74, (6) 196-203.

Mertens, H. W. 1990, *Evaluation of functional color vision requirements and current color vision screening tests for air traffic control specialists*, Office of Aviation Medicine, Federal Aviation Administration, Washington D.C., DOT/FAA/AM-90/9.

Mertens, H.W. & Milburn, N.J. 1996. Performance of color-dependent air traffic control tasks as a function of color vision deficiency. *Aviat.Space Environ.Med.*, 67, (10) 919-927.

National Research Council-National Academy of Sciences, C. o. V. 1981, *Procedures for testing color vision*, National Academy Press, Washington, DC, Report of working group 41.

Nothdurft, H.C. 1993. The role of features in preattentive vision: comparison of orientation, motion and color cues. *Vision Res.*, 33, (14) 1937-1958.

O'Brien, K.A., Cole, B.L., Maddocks, J.D., & Forbes, A.B. 2002. Color and defective color vision as factors in the conspicuity of signs and signals. *Hum.Factors*, 44, (4) 665-675.

Pinker, S. 1984. Visual cognition: an introduction. *Cognition*, 18, (1-3) 1-63.

Ramaswamy, S. & Hovis, J.K. 2004. Ability of the D-15 panel tests and HRR pseudoisochromatic plates to predict performance in naming VDT colors. *Vis.Neurosci*, 21, (3) 455-460.

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 Colour Association, Granada (Granada, 2005)* 979-982

Rodriguez-Carmona, M., O'Neill-Biba, M., & Barbur, J.L. 2012. Assessing the severity of color vision loss with implications for aviation and other occupational environments. *Aviat.Space Environ.Med.*, 83, (1) 19-29.

Sachtler, W.L. & Zaidi, Q. 1992. Chromatic and luminance signals in visual memory. *J Opt.Soc.Am.A*, 9, (6) 877-894.

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*, K. R. Gegenfurtner & L. T. Sharpe, eds., Cambridge: Cambridge University Press, pp. 3-52.

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, (2) 229-239.

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

Treisman, A.M. & Gelade, G. 1980. A feature-integration theory of attention. *Cogn Psychol.*, 12, (1) 97-136.

U.S.Department of Transportation - Federal Aviation Administration 2007, *HF-STD-002*.

Walkey, H.C., Barbur, J.L., Harlow, J.A., Hurden, A., Moorhead, I.R., & Taylor, J.A. 2005. Effective contrast of colored stimuli in the mesopic range: a metric for perceived contrast based on achromatic luminance contrast. *J Opt.Soc.Am.A Opt.Image Sci.Vis.*, 22, (1) 17-28.

Walkey, H.C., Harlow, J.A., & Barbur, J.L. 2006. Characterising mesopic spectral sensitivity from reaction times. *Vision Res*, 46, (25) 4232-4243.

Wright, W.D. 1946. *Researches on Normal and Defective Colour Vision* London, Henry Kimpton.

Xing, J. & Schroeder, D. J. 2006, *Reexamination of Color Vision Standards, Part I: Status of Color ATC Displays and Demography of Color-Deficit Controllers*, FAA Technical Report, Washington, DC: Federal Aviation Administration, DOT/FAA/AM-06/02.

Yamagishi, N. & Melara, R.D. 2001. Informational primacy of visual dimensions: specialized roles for luminance and chromaticity in figure-ground perception. *Percept.Psychophys.*, 63, (5) 824-846.